GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 29, 2005, 01:45:27 ; Search time 5154 Seconds
(without alignments)
10445.042 Million cell updates/sec

Title: US-09-979-549-2
Perfect score: 1111
Sequence: 1 gaagtcagaaggccgttcag.....ccgagaagagagagagagag 1111

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1111	100.0	1111	6	BD263890	BD263890 Promoter
	2	1111	100.0	1111	6	AX047682	AX047682 Sequence
	3	1111	100.0	2687	6	BD263889	BD263889 Promoter
	4	1111	100.0	2687	6	AX047681	AX047681 Sequence
	5	54	4.9	596	8	TAE9762	AJ009762 Triticum
	6	32	2.9	629	8	TAE404845	AJ404845 Triticum
c	7	26	2.3	228676	2	BX901914	BX901914 Danio rer
c	8	24	2.2	146015	2	AC021823	AC021823 Homo sapi
	9	24	2.2	152251	9	AC092642	AC092642 Homo sapi
	10	24	2.2	155054	2	BX927299	BX927299 Danio rer
c	11	24	2.2	160804	5	BX001022	BX001022 Zebrafish
c	12	24	2.2	185217	2	BX927400	BX927400 Danio rer
c	13	24	2.2	247387	5	BX649502	BX649502 Zebrafish
c	14	24	2.2	263169	2	BX927111	BX927111 Danio rer
c	15	24	2.2	291544	2	AC128496	AC128496 Rattus no
	16	23	2.1	116090	10	AL772222	AL772222 Mouse DNA
c	17	23	2.1	163315	2	AC142056	AC142056 Rattus no
c	18	23	2.1	166125	2	AC019327	AC019327 Homo sapi
	19	23	2.1	177332	9	AC027290	AC027290 Homo sapi
c	20	23	2.1	187745	2	AC147781	AC147781 Mus muscu
	21	23	2.1	188285	9	AL135842	AL135842 Human DNA
	22	22	2.0	141	6	CQ466652	CQ466652 Sequence
	23	22	2.0	144	6	CQ466387	CQ466387 Sequence
	24	22	2.0	154	6	CQ466521	CQ466521 Sequence
	25	22	2.0	213	6	CQ467148	CQ467148 Sequence
c	26	22	2.0	249	6	AR554285	AR554285 Sequence
c	27	22	2.0	1074	6	CQ413259	CQ413259 Sequence
c	28	22	2.0	12634	8	AF210616	AF210616 Zea mays
c	29	22	2.0	22270	8	AY702552	AY702552 Zea mays
c	30	22	2.0	24935	6	CQ585068	CQ585068 Sequence
	31	22	2.0	41305	2	AC014956	AC014956 Drosophil
c	32	22	2.0	85305	5	AL929231	AL929231 Zebrafish
c	33	22	2.0	126552	9	AC006431	AC006431 Homo sapi
	34	22	2.0	131634	2	AC147655	AC147655 Pan trogl
c	35	22	2.0	131634	2	AC147655	AC147655 Pan trogl
c	36	22	2.0	142275	2	CR788316	CR788316 Danio rer
	37	22	2.0	142294	2	AP002352	AP002352 Homo sapi
c	38	22	2.0	145417	2	AC138169	AC138169 Sus scrof
	39	22	2.0	146106	9	AC147342	AC147342 Pan trogl
	40	22	2.0	148076	2	AC151848	AC151848 Pan trogl
c	41	22	2.0	148811	9	AC073857	AC073857 Homo sapi
	42	22	2.0	153440	3	AC105263	AC105263 Drosophil
c	43	22	2.0	155363	5	BX088582	BX088582 Zebrafish
	44	22	2.0	160237	2	AC149093	AC149093 Pan trogl
c	45	22	2.0	160602	9	AC147661	AC147661 Pan trogl

ALIGNMENTS

RESULT 1
BD263890
LOCUS BD263890 1111 bp DNA linear PAT 17-JUL-2003
DEFINITION Promoter of thioredoxine TaTrxh2 in wheat.
ACCESSION BD263890
VERSION BD263890.1 GI:33073658
KEYWORDS JP 2002543844-A/2.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 1111)
AUTHORS Gautier,M.F., Ihorai,T. and Joudrier,P.
TITLE Promoter of thioredoxine TaTrxh2 in wheat
JOURNAL Patent: JP 2002543844-A 2 24-DEC-2002;
INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE
COMMENT OS Triticum aestivum (common wheat)
PN JP 2002543844-A/2
PD 24-DEC-2002
PF 17-MAY-2000 JP 2000618471
PR 17-MAY-1999 FR 99/06231
PI MARIE FRANCOISE GAUTIER,TANIA IHORAI,PHILIPPE JOUDRIER PC C12N15/09,A01H5/00,C12N5/10//(C12N5/10,C12R1:91),C12N15/00, PC C12N5/00,
PC (C12N5/00,C12R1:91)
CC Promoter of thioredoxine TaTrxh2 in wheat
FH Key Location/Qualifiers
FT source 1. .1111
FT /organism='Triticum aestivum (common wheat)'.
FEATURES Location/Qualifiers
source 1. .1111
/organism="Triticum aestivum"
/mol_type="genomic DNA"
/db_xref="taxon:4565"
ORIGIN

Query Match 100.0%; Score 1111; DB 6; Length 1111;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAGTCAGAAGGCCGTTCAGAATTGTTGGAGGACTCGAAAAAAAGAAGGGGAGCCCAGGC 60

```
Db          1 GAAGTCAGAAGGCCGTTCAGAATTGTTGGAGGACTCGAAAAAAAGAAGGGGAGCCCAGGC 60
Qy         61 AGACGACGGGGCGGCATGTGCCTGTTCCTTGGCGAGGCGTCTAGCTTTGGCAGCCGCCGC 120
Db         61 AGACGACGGGGCGGCATGTGCCTGTTCCTTGGCGAGGCGTCTAGCTTTGGCAGCCGCCGC 120
Qy        121 CGCTTTTCTCCTTGGGTGGGCGCGCGAGCTCCCCGAGTTTGAGCCGCAATTTTTTTACAT 180
Db        121 CGCTTTTCTCCTTGGGTGGGCGCGCGAGCTCCCCGAGTTTGAGCCGCAATTTTTTTACAT 180
Qy        181 TTTATGGCGATGGCGTCAGGCGTTTATCTAGGCGTCTGGGAGGGTACATTTGAAGATGTG 240
Db        181 TTTATGGCGATGGCGTCAGGCGTTTATCTAGGCGTCTGGGAGGGTACATTTGAAGATGTG 240
Qy        241 CCACCAACTCCAAACCGACAACCCTGTATCTGAGCATGCCTCATGCCTCTCCTTCATGCC 300
Db        241 CCACCAACTCCAAACCGACAACCCTGTATCTGAGCATGCCTCATGCCTCTCCTTCATGCC 300
Qy        301 TCCCTTTGGGTGAGGTCATGTGCCCTTGGCGGCGAGTGGCTTCCCGTTTAGAGCAAGTAT 360
Db        301 TCCCTTTGGGTGAGGTCATGTGCCCTTGGCGGCGAGTGGCTTCCCGTTTAGAGCAAGTAT 360
Qy        361 AATAAGTCCTAGTCAGCTGGCTATAAGATGTTCCACATCAGCAAATCCTTAAACTGGAGG 420
Db        361 AATAAGTCCTAGTCAGCTGGCTATAAGATGTTCCACATCAGCAAATCCTTAAACTGGAGG 420
Qy        421 AGAAAGAAAGTAGGAGTGAGAAGGGCGTCGGCGCTTCGTCAATCGCTAGCGATAGCACAA 480
Db        421 AGAAAGAAAGTAGGAGTGAGAAGGGCGTCGGCGCTTCGTCAATCGCTAGCGATAGCACAA 480
Qy        481 GCTCCCATGGAATCGAGCCAACATGCAACCCGCACAATGACTAAAGGCAAACGCCAGCCA 540
Db        481 GCTCCCATGGAATCGAGCCAACATGCAACCCGCACAATGACTAAAGGCAAACGCCAGCCA 540
Qy        541 ATCAGTATGCCTTTCTCTGCATCTTTCTTCATGCAAGCATTAAATACTATAGCTAATCTA 600
Db        541 ATCAGTATGCCTTTCTCTGCATCTTTCTTCATGCAAGCATTAAATACTATAGCTAATCTA 600
Qy        601 CAGCCAGTTTATTATATAAACAGGCTATATAGCTGACCTGGCAGTGCTATAGAGCCGGCA 660
Db        601 CAGCCAGTTTATTATATAAACAGGCTATATAGCTGACCTGGCAGTGCTATAGAGCCGGCA 660
Qy        661 GCCGGCTCTTCTATTAGCTTTGCTCTTATGGCTACATCTGTGTGAGCAGTCGATTGATTC 720
Db        661 GCCGGCTCTTCTATTAGCTTTGCTCTTATGGCTACATCTGTGTGAGCAGTCGATTGATTC 720
Qy        721 AAACAACAAATCCGGGCGTTCAGCAAGTCGGAATGAATTTCGGCTCATCACTCATTGTCG 780
Db        721 AAACAACAAATCCGGGCGTTCAGCAAGTCGGAATGAATTTCGGCTCATCACTCATTGTCG 780
Qy        781 TGGGCCTCACGCGTATTCGCCTAACCGTGTTTGAATCAGACCCTCACGAAGCCACGGCTC 840
Db        781 TGGGCCTCACGCGTATTCGCCTAACCGTGTTTGAATCAGACCCTCACGAAGCCACGGCTC 840
Qy        841 CAGCGACCCGTTCACCACGTCAGCCTAAAAAAAGAAAAAAAAACTGTTCAATCACACGCC 900
Db        841 CAGCGACCCGTTCACCACGTCAGCCTAAAAAAAGAAAAAAAAACTGTTCAATCACACGCC 900
Qy        901 CATCTGAACCGTTCAACAGCCCCACGTAATTTCGCGCACCAGCAAAGGGCATATCCGTCA 960
Db        901 CATCTGAACCGTTCAACAGCCCCACGTAATTTCGCGCACCAGCAAAGGGCATATCCGTCA 960
Qy        961 TAGCGAGCGCATAAATTCTGATTCCTGCCTGCCTGCCGGACAATTTATCTTTGGGGAGGC 1020
Db        961 TAGCGAGCGCATAAATTCTGATTCCTGCCTGCCTGCCGGACAATTTATCTTTGGGGAGGC 1020
Qy       1021 GGGCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAATCAACA 1080
Db       1021 GGGCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAATCAACA 1080
Qy       1081 AGCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 1111
```

```
Db       1081 AGCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 1111
```

```
RESULT 2
AX047682
LOCUS       AX047682               1111 bp    DNA     linear   PAT 15-DEC-2000
DEFINITION  Sequence 2 from Patent WO0070065.
ACCESSION   AX047682
VERSION     AX047682.1  GI:11876717
KEYWORDS    .
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1
  AUTHORS   Gautier,M.F., Ihorai,T. and Joudrier,P.
  TITLE     Promoter of thioredoxine tatrxh2 in wheat
  JOURNAL   Patent: WO 0070065-A 2 23-NOV-2000;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR)
FEATURES             Location/Qualifiers
     source          1. .1111
                     /organism="Triticum aestivum"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4565"
ORIGIN

  Query Match             100.0%;  Score 1111;  DB 6;  Length 1111;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1111;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAAGTCAGAAGGCCGTTCAGAATTGTTGGAGGACTCGAAAAAAAGAAGGGGAGCCCAGGC 60
Db          1 GAAGTCAGAAGGCCGTTCAGAATTGTTGGAGGACTCGAAAAAAAGAAGGGGAGCCCAGGC 60
Qy         61 AGACGACGGGGCGGCATGTGCCTGTTCCTTGGCGAGGCGTCTAGCTTTGGCAGCCGCCGC 120
Db         61 AGACGACGGGGCGGCATGTGCCTGTTCCTTGGCGAGGCGTCTAGCTTTGGCAGCCGCCGC 120
Qy        121 CGCTTTTCTCCTTGGGTGGGCGCGCGAGCTCCCCGAGTTTGAGCCGCAATTTTTTTACAT 180
Db        121 CGCTTTTCTCCTTGGGTGGGCGCGCGAGCTCCCCGAGTTTGAGCCGCAATTTTTTTACAT 180
Qy        181 TTTATGGCGATGGCGTCAGGCGTTTATCTAGGCGTCTGGGAGGGTACATTTGAAGATGTG 240
Db        181 TTTATGGCGATGGCGTCAGGCGTTTATCTAGGCGTCTGGGAGGGTACATTTGAAGATGTG 240
Qy        241 CCACCAACTCCAAACCGACAACCCTGTATCTGAGCATGCCTCATGCCTCTCCTTCATGCC 300
Db        241 CCACCAACTCCAAACCGACAACCCTGTATCTGAGCATGCCTCATGCCTCTCCTTCATGCC 300
Qy        301 TCCCTTTGGGTGAGGTCATGTGCCCTTGGCGGCGAGTGGCTTCCCGTTTAGAGCAAGTAT 360
Db        301 TCCCTTTGGGTGAGGTCATGTGCCCTTGGCGGCGAGTGGCTTCCCGTTTAGAGCAAGTAT 360
Qy        361 AATAAGTCCTAGTCAGCTGGCTATAAGATGTTCCACATCAGCAAATCCTTAAACTGGAGG 420
Db        361 AATAAGTCCTAGTCAGCTGGCTATAAGATGTTCCACATCAGCAAATCCTTAAACTGGAGG 420
Qy        421 AGAAAGAAAGTAGGAGTGAGAAGGGCGTCGGCGCTTCGTCAATCGCTAGCGATAGCACAA 480
Db        421 AGAAAGAAAGTAGGAGTGAGAAGGGCGTCGGCGCTTCGTCAATCGCTAGCGATAGCACAA 480
Qy        481 GCTCCCATGGAATCGAGCCAACATGCAACCCGCACAATGACTAAAGGCAAACGCCAGCCA 540
Db        481 GCTCCCATGGAATCGAGCCAACATGCAACCCGCACAATGACTAAAGGCAAACGCCAGCCA 540
Qy        541 ATCAGTATGCCTTTCTCTGCATCTTTCTTCATGCAAGCATTAAATACTATAGCTAATCTA 600
Db        541 ATCAGTATGCCTTTCTCTGCATCTTTCTTCATGCAAGCATTAAATACTATAGCTAATCTA 600
Qy        601 CAGCCAGTTTATTATATAAACAGGCTATATAGCTGACCTGGCAGTGCTATAGAGCCGGCA 660
```

```
Db        601 CAGCCAGTTTATTATATAAACAGGCTATATAGCTGACCTGGCAGTGCTATAGAGCCGGCA 660

Qy        661 GCCGGCTCTTCTATTAGCTTTGCTCTTATGGCTACATCTGTGTGAGCAGTCGATTGATTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCGGCTCTTCTATTAGCTTTGCTCTTATGGCTACATCTGTGTGAGCAGTCGATTGATTC 720

Qy        721 AAACAACAAATCCGGGCGTTCAGCAAGTCGGAATGAATTTCGGCTCATCACTCATTGTCG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAACAACAAATCCGGGCGTTCAGCAAGTCGGAATGAATTTCGGCTCATCACTCATTGTCG 780

Qy        781 TGGGCCTCACGCGTATTCGCCTAACCGTGTTTGAATCAGACCCTCACGAAGCCACGGCTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGGGCCTCACGCGTATTCGCCTAACCGTGTTTGAATCAGACCCTCACGAAGCCACGGCTC 840

Qy        841 CAGCGACCCGTTCACCACGTCAGCCTAAAAAAAGAAAAAAAAACTGTTCAATCACACGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CAGCGACCCGTTCACCACGTCAGCCTAAAAAAAGAAAAAAAAACTGTTCAATCACACGCC 900

Qy        901 CATCTGAACCGTTCAACAGCCCCACGTAATTTCGCGCACCAGCAAAGGGCATATCCGTCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CATCTGAACCGTTCAACAGCCCCACGTAATTTCGCGCACCAGCAAAGGGCATATCCGTCA 960

Qy        961 TAGCGAGCGCATAAATTCTGATTCCTGCCTGCCTGCCGGACAATTTATCTTTGGGGAGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TAGCGAGCGCATAAATTCTGATTCCTGCCTGCCTGCCGGACAATTTATCTTTGGGGAGGC 1020

Qy       1021 GGGCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAATCAACA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGGCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAATCAACA 1080

Qy       1081 AGCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 1111
              ||||||||||||||||||||||||||||||||
Db       1081 AGCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 1111
```

```
RESULT 3
BD263889
LOCUS       BD263889               2687 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Promoter of thioredoxine TaTrxh2 in wheat.
ACCESSION   BD263889
VERSION     BD263889.1  GI:33073657
KEYWORDS    JP 2002543844-A/1.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 2687)
  AUTHORS   Gautier,M.F., Ihorai,T. and Joudrier,P.
  TITLE     Promoter of thioredoxine TaTrxh2 in wheat
  JOURNAL   Patent: JP 2002543844-A 1 24-DEC-2002;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE
COMMENT     OS   Triticum aestivum (common wheat)
            PN   JP 2002543844-A/1
            PD   24-DEC-2002
            PF   17-MAY-2000 JP 2000618471
            PR   17-MAY-1999 FR     99/06231
            PI   MARIE FRANCOISE GAUTIER,TANIA IHORAI,PHILIPPE JOUDRIER PC
            C12N15/09,A01H5/00,C12N5/10//(C12N5/10,C12R1:91),C12N15/00, PC
            C12N5/00,
            PC   (C12N5/00,C12R1:91)
            CC   Promoter of thioredoxine TaTrxh2 in wheat
            FH   Key             Location/Qualifiers
            FT   exon            (1112). .(1231)
            FT   intron          (1232). .(2203)
            FT   exon            (2204). .(2326)
            FT   intron          (2327). .(2420)
            FT   exon            (2421). .(2558)
            FT   CDS             (1112). .(1231)
            FT   CDS             (2204). .(2326)
            FT   CDS             (2421). .(2558).
FEATURES             Location/Qualifiers
     source          1. .2687
                     /organism="Triticum aestivum"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4565"
ORIGIN
```

```
  Query Match           100.0%;  Score 1111;  DB 6;  Length 2687;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1111;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAAGTCAGAAGGCCGTTCAGAATTGTTGGAGGACTCGAAAAAAAAGAAGGGGAGCCCAGGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAGTCAGAAGGCCGTTCAGAATTGTTGGAGGACTCGAAAAAAAAGAAGGGGAGCCCAGGC 60

Qy         61 AGACGACGGGGCGGCATGTGCCTGTTCCTTGGCGAGGCGTCTAGCTTTGGCAGCCGCCGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGACGACGGGGCGGCATGTGCCTGTTCCTTGGCGAGGCGTCTAGCTTTGGCAGCCGCCGC 120

Qy        121 CGCTTTTCTCCTTGGGTGGGCGCGCGAGCTCCCCGAGTTTGAGCCGCAATTTTTTTACAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCTTTTCTCCTTGGGTGGGCGCGCGAGCTCCCCGAGTTTGAGCCGCAATTTTTTTACAT 180

Qy        181 TTTATGGCGATGGCGTCAGGCGTTTATCTAGGCGTCTGGGAGGGTACATTTGAAGATGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTTATGGCGATGGCGTCAGGCGTTTATCTAGGCGTCTGGGAGGGTACATTTGAAGATGTG 240

Qy        241 CCACCAACTCCAAACCGACAACCCTGTATCTGAGCATGCCTCATGCCTCTCCTTCATGCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCACCAACTCCAAACCGACAACCCTGTATCTGAGCATGCCTCATGCCTCTCCTTCATGCC 300

Qy        301 TCCCTTTGGGTGAGGTCATGTGCCCTTGGCGGCGAGTGGCTTCCCGTTTAGAGCAAGTAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCCCTTTGGGTGAGGTCATGTGCCCTTGGCGGCGAGTGGCTTCCCGTTTAGAGCAAGTAT 360

Qy        361 AATAAGTCCTAGTCAGCTGGCTATAAGATGTTCCACATCAGCAAATCCTTAAACTGGAGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AATAAGTCCTAGTCAGCTGGCTATAAGATGTTCCACATCAGCAAATCCTTAAACTGGAGG 420

Qy        421 AGAAAGAAAGTAGGAGTGAGAAGGGCGTCGGCGCTTCGTCAATCGCTAGCGATAGCACAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGAAAGAAAGTAGGAGTGAGAAGGGCGTCGGCGCTTCGTCAATCGCTAGCGATAGCACAA 480

Qy        481 GCTCCCATGGAATCGAGCCAACATGCAACCCGCACAATGACTAAAGGCAAACGCCAGCCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTCCCATGGAATCGAGCCAACATGCAACCCGCACAATGACTAAAGGCAAACGCCAGCCA 540

Qy        541 ATCAGTATGCCTTTCTCTGCATCTTTCTTCATGCAAGCATTAAATACTATAGCTAATCTA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATCAGTATGCCTTTCTCTGCATCTTTCTTCATGCAAGCATTAAATACTATAGCTAATCTA 600

Qy        601 CAGCCAGTTTATTATATAAACAGGCTATATAGCTGACCTGGCAGTGCTATAGAGCCGGCA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAGCCAGTTTATTATATAAACAGGCTATATAGCTGACCTGGCAGTGCTATAGAGCCGGCA 660

Qy        661 GCCGGCTCTTCTATTAGCTTTGCTCTTATGGCTACATCTGTGTGAGCAGTCGATTGATTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCGGCTCTTCTATTAGCTTTGCTCTTATGGCTACATCTGTGTGAGCAGTCGATTGATTC 720

Qy        721 AAACAACAAATCCGGGCGTTCAGCAAGTCGGAATGAATTTCGGCTCATCACTCATTGTCG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAACAACAAATCCGGGCGTTCAGCAAGTCGGAATGAATTTCGGCTCATCACTCATTGTCG 780

Qy        781 TGGGCCTCACGCGTATTCGCCTAACCGTGTTTGAATCAGACCCTCACGAAGCCACGGCTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGGGCCTCACGCGTATTCGCCTAACCGTGTTTGAATCAGACCCTCACGAAGCCACGGCTC 840

Qy        841 CAGCGACCCGTTCACCACGTCAGCCTAAAAAAAGAAAAAAAAACTGTTCAATCACACGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CAGCGACCCGTTCACCACGTCAGCCTAAAAAAAGAAAAAAAAACTGTTCAATCACACGCC 900

Qy        901 CATCTGAACCGTTCAACAGCCCCACGTAATTTCGCGCACCAGCAAAGGGCATATCCGTCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CATCTGAACCGTTCAACAGCCCCACGTAATTTCGCGCACCAGCAAAGGGCATATCCGTCA 960
```

```
Qy        961 TAGCGAGCGCATAAATTCTGATTCCTGCCTGCCTGCCGGACAATTTATCTTTGGGGAGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TAGCGAGCGCATAAATTCTGATTCCTGCCTGCCTGCCGGACAATTTATCTTTGGGGAGGC 1020

Qy       1021 GGGCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAATCAACA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGGCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAATCAACA 1080

Qy       1081 AGCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 1111
              ||||||||||||||||||||||||||||||||
Db       1081 AGCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 1111


RESULT 4
AX047681
LOCUS       AX047681                2687 bp    DNA     linear   PAT 15-DEC-2000
DEFINITION  Sequence 1 from Patent WO0070065.
ACCESSION   AX047681
VERSION     AX047681.1  GI:11876716
KEYWORDS    .
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1
  AUTHORS   Gautier,M.F., Ihorai,T. and Joudrier,P.
  TITLE     Promoter of thioredoxine tatrxh2 in wheat
  JOURNAL   Patent: WO 0070065-A 1 23-NOV-2000;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR)
FEATURES             Location/Qualifiers
     source          1. .2687
                     /organism="Triticum aestivum"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:4565"
     exon            1112. .1231
     intron          1232. .2203
     exon            2204. .2326
     intron          2327. .2420
     exon            2421. .2558
ORIGIN

  Query Match            100.0%;  Score 1111;  DB 6;  Length 2687;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1111;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAAGTCAGAAGGCCGTTCAGAATTGTTGGAGGACTCGAAAAAAAGAAGGGGAGCCCAGGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAGTCAGAAGGCCGTTCAGAATTGTTGGAGGACTCGAAAAAAAGAAGGGGAGCCCAGGC 60

Qy         61 AGACGACGGGGCGGCATGTGCCTGTTCCTTGGCGAGGCGTCTAGCTTTGGCAGCCGCCGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGACGACGGGGCGGCATGTGCCTGTTCCTTGGCGAGGCGTCTAGCTTTGGCAGCCGCCGC 120

Qy        121 CGCTTTTCTCCTTGGGTGGGCGCGCGAGCTCCCCGAGTTTGAGCCGCAATTTTTTTACAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCTTTTCTCCTTGGGTGGGCGCGCGAGCTCCCCGAGTTTGAGCCGCAATTTTTTTACAT 180

Qy        181 TTTATGGCGATGGCGTCAGGCGTTTATCTAGGCGTCTGGGAGGGTACATTTGAAGATGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTTATGGCGATGGCGTCAGGCGTTTATCTAGGCGTCTGGGAGGGTACATTTGAAGATGTG 240

Qy        241 CCACCAACTCCAAACCGACAACCCTGTATCTGAGCATGCCTCATGCCTCTCCTTCATGCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCACCAACTCCAAACCGACAACCCTGTATCTGAGCATGCCTCATGCCTCTCCTTCATGCC 300

Qy        301 TCCCTTTGGGTGAGGTCATGTGCCCTTGGCGGCGAGTGGCTTCCCGTTTAGAGCAAGTAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCCCTTTGGGTGAGGTCATGTGCCCTTGGCGGCGAGTGGCTTCCCGTTTAGAGCAAGTAT 360

Qy        361 AATAAGTCCTAGTCAGCTGGCTATAAGATGTTCCACATCAGCAAATCCTTAAACTGGAGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AATAAGTCCTAGTCAGCTGGCTATAAGATGTTCCACATCAGCAAATCCTTAAACTGGAGG 420

Qy        421 AGAAAGAAAGTAGGAGTGAGAAGGGCGTCGGCGCTTCGTCAATCGCTAGCGATAGCACAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGAAAGAAAGTAGGAGTGAGAAGGGCGTCGGCGCTTCGTCAATCGCTAGCGATAGCACAA 480

Qy        481 GCTCCCATGGAATCGAGCCAACATGCAACCCGCACAATGACTAAAGGCAAACGCCAGCCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTCCCATGGAATCGAGCCAACATGCAACCCGCACAATGACTAAAGGCAAACGCCAGCCA 540

Qy        541 ATCAGTATGCCTTTCTCTGCATCTTTCTTCATGCAAGCATTAAATACTATAGCTAATCTA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATCAGTATGCCTTTCTCTGCATCTTTCTTCATGCAAGCATTAAATACTATAGCTAATCTA 600

Qy        601 CAGCCAGTTTATTATATAAACAGGCTATATAGCTGACCTGGCAGTGCTATAGAGCCGGCA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAGCCAGTTTATTATATAAACAGGCTATATAGCTGACCTGGCAGTGCTATAGAGCCGGCA 660

Qy        661 GCCGGCTCTTCTATTAGCTTTGCTCTTATGGCTACATCTGTGTGAGCAGTCGATTGATTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCGGCTCTTCTATTAGCTTTGCTCTTATGGCTACATCTGTGTGAGCAGTCGATTGATTC 720

Qy        721 AAACAACAAATCCGGGCGTTCAGCAAGTCGGAATGAATTTCGGCTCATCACTCATTGTCG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAACAACAAATCCGGGCGTTCAGCAAGTCGGAATGAATTTCGGCTCATCACTCATTGTCG 780

Qy        781 TGGGCCTCACGCGTATTCGCCTAACCGTGTTTGAATCAGACCCTCACGAAGCCACGGCTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGGGCCTCACGCGTATTCGCCTAACCGTGTTTGAATCAGACCCTCACGAAGCCACGGCTC 840

Qy        841 CAGCGACCCGTTCACCACGTCAGCCTAAAAAAAGAAAAAAAAACTGTTCAATCACACGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CAGCGACCCGTTCACCACGTCAGCCTAAAAAAAGAAAAAAAAACTGTTCAATCACACGCC 900

Qy        901 CATCTGAACCGTTCAACAGCCCCACGTAATTTCGCGCACCAGCAAAGGGCATATCCGTCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CATCTGAACCGTTCAACAGCCCCACGTAATTTCGCGCACCAGCAAAGGGCATATCCGTCA 960

Qy        961 TAGCGAGCGCATAAATTCTGATTCCTGCCTGCCTGCCGGACAATTTATCTTTGGGGAGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TAGCGAGCGCATAAATTCTGATTCCTGCCTGCCTGCCGGACAATTTATCTTTGGGGAGGC 1020

Qy       1021 GGGCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAATCAACA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGGCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAATCAACA 1080

Qy       1081 AGCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 1111
              ||||||||||||||||||||||||||||||||
Db       1081 AGCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 1111


RESULT 5
TAE9762
LOCUS       TAE9762                 596 bp    mRNA    linear   PLN 15-AUG-2001
DEFINITION  Triticum aestivum mRNA for thioredoxin H.
ACCESSION   AJ009762
VERSION     AJ009762.1  GI:4138593
KEYWORDS    thioredoxin H.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1
  AUTHORS   Serrato,A.J., Crespo,J.L., Florencio,F.J. and Cejudo,F.J.
  TITLE     Characterization of two thioredoxins h with predominant
            localization in the nucleus of aleurone and scutellum cells of
            germinating wheat seeds
  JOURNAL   Plant Mol. Biol. 46 (3), 361-371 (2001)
  MEDLINE   21380673
   PUBMED   11488482
REFERENCE   2  (bases 1 to 596)
  AUTHORS   Cejudo,F.J.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (24-JUL-1998) Cejudo F.J., Instituto de Bioquimica
            Vegetal Y Fotosintesis, Universidad de Sevilla Y CSIC, Avda Americo
            Vespucio s/n, 41092-Sevilla, SPAIN
FEATURES             Location/Qualifiers
     source          1. .596
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
     gene            1. .596
                     /gene="thioredoxin H"
     3'UTR           436. .596
                     /gene="thioredoxin H"
     polyA_signal    462. .466
                     /gene="thioredoxin H"
ORIGIN

  Query Match             4.9%;  Score 54;  DB 8;  Length 596;
  Best Local Similarity   100.0%;  Pred. No. 2e-18;
  Matches   54;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1058 AAAGTGCGCGTGAGAAATCAACAAGCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 1111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAAGTGCGCGTGAGAAATCAACAAGCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 54


RESULT 6
TAE404845
LOCUS       TAE404845               629 bp    mRNA    linear   PLN 15-AUG-2001
DEFINITION  Triticum aestivum mRNA for thioredoxin h.
ACCESSION   AJ404845
VERSION     AJ404845.1  GI:8980490
KEYWORDS    thioredoxin H.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1
  AUTHORS   Serrato,A.J., Crespo,J.L., Florencio,F.J. and Cejudo,F.J.
  TITLE     Characterization of two thioredoxins h with predominant
            localization in the nucleus of aleurone and scutellum cells of
            germinating wheat seeds
  JOURNAL   Plant Mol. Biol. 46 (3), 361-371 (2001)
  MEDLINE   21380673
   PUBMED   11488482
REFERENCE   2  (bases 1 to 629)
  AUTHORS   Cejudo,F.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUL-2000) Cejudo F.J., Instituto de Bioquimica
            Vegetal y Fotosintesis, Universidad de Sevilla, Avda Americo
            Vespucio s/n, 41092-Sevilla, SPAIN
FEATURES             Location/Qualifiers
     source          1. .629
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /tissue_type="aleurone"
                     /dev_stage="germinating seed"
                     /country="Spain"
     5'UTR           1. .111
     CDS             112. .489
                     /note="ORF"
                     /codon_start=1
                     /product="thioredoxin h"
                     /protein_id="CAB96931.1"
                     /db_xref="GI:8980491"
                     /db_xref="GOA:Q9LDX4"
                     /db_xref="UniProt/TrEMBL:Q9LDX4"
                     /translation="MAASAATATAAAVGAGEVISVHSLEQWTMQIEEANAAKKLVVID
                     FTASWCGPCRIMAPIFADLAKKFPAAVFLKVDVDELKSIAEQFSVEAMPTFLFMKEGD
                     VKDRVVGAIKEELTNKVGLHAAQ"
     3'UTR           490. .629
     polyA_signal    516. .523
ORIGIN

  Query Match             2.9%;  Score 32;  DB 8;  Length 629;
  Best Local Similarity   100.0%;  Pred. No. 6.6e-06;
  Matches   32;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1080 AAGCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 1111
              ||||||||||||||||||||||||||||||||
Db         58 AAGCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 89


RESULT 7
BX901914/c
LOCUS       BX901914               228676 bp    DNA     linear   HTG 10-OCT-2004
DEFINITION  Danio rerio clone DKEY-27F18, WORKING DRAFT SEQUENCE, 17 unordered
            pieces.
ACCESSION   BX901914
VERSION     BX901914.6  GI:54021834
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 228676)
  AUTHORS   McLaren,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-OCT-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Oct 10, 2004 this sequence version replaced gi:46194670.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -------------- Project Information
            Center project name: zK27F18
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 222727 bases at least Q40
            Consensus quality: 223391 bases at least Q30
            Consensus quality: 224176 bases at least Q20
            Insert size: 227076; sum-of-contigs
            Insert size: 222615; 5.7% error; agarose-fp
            Quality coverage: 6.94x in Q20 bases; sum-of-contigs Quality
            coverage: 7.21x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 17 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    10555: contig of 10555 bp in length
            *    10556    10655: gap of 100 bp
            *    10656    20734: contig of 10079 bp in length
            *    20735    20834: gap of 100 bp
            *    20835    29431: contig of 8597 bp in length
            *    29432    29531: gap of 100 bp
            *    29532    47378: contig of 17847 bp in length
            *    47379    47478: gap of 100 bp
            *    47479    49583: contig of 2105 bp in length
            *    49584    49683: gap of 100 bp
            *    49684    75283: contig of 25600 bp in length
            *    75284    75383: gap of 100 bp
            *    75384    93589: contig of 18206 bp in length
            *    93590    93689: gap of 100 bp
```

```
            *    93690    96129: contig of 2440 bp in length
            *    96130    96229: gap of 100 bp
            *    96230   109725: contig of 13496 bp in length
            *   109726   109825: gap of 100 bp
            *   109826   124219: contig of 14394 bp in length
            *   124220   124319: gap of 100 bp
            *   124320   129512: contig of 5193 bp in length
            *   129513   129612: gap of 100 bp
            *   129613   134036: contig of 4424 bp in length
            *   134037   134136: gap of 100 bp
            *   134137   151328: contig of 17192 bp in length
            *   151329   151428: gap of 100 bp
            *   151429   167242: contig of 15814 bp in length
            *   167243   167342: gap of 100 bp
            *   167343   179999: contig of 12657 bp in length
            *   180000   180099: gap of 100 bp
            *   180100   190992: contig of 10893 bp in length
            *   190993   191092: gap of 100 bp
            *   191093   228676: contig of 37584 bp in length.
FEATURES             Location/Qualifiers
     source          1. .228676
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-27F18"
                     /clone_lib="DanioKey"
     misc_feature    1. .10555
                     /note="assembly_fragment:00271
                     fragment_chain:1"
     misc_feature    10656. .20734
                     /note="assembly_fragment:00142
                     fragment_chain:1"
     misc_feature    20835. .29431
                     /note="assembly_fragment:00056
                     fragment_chain:1"
     misc_feature    29532. .47378
                     /note="assembly_fragment:01109
                     fragment_chain:1"
     misc_feature    47479. .49583
                     /note="assembly_fragment:02902
                     fragment_chain:1"
     misc_feature    49684. .75283
                     /note="assembly_fragment:02010
                     fragment_chain:2"
     misc_feature    75384. .93589
                     /note="assembly_fragment:01546
                     fragment_chain:2"
     misc_feature    93690. .96129
                     /note="assembly_fragment:02882
                     fragment_chain:2"
     misc_feature    96230. .109725
                     /note="assembly_fragment:00733
                     fragment_chain:2"
     misc_feature    109826. .124219
                     /note="assembly_fragment:00914
                     fragment_chain:2"
     misc_feature    124320. .129512
                     /note="assembly_fragment:02928
                     fragment_chain:3"
     misc_feature    129613. .134036
                     /note="assembly_fragment:00001
                     fragment_chain:3"
     misc_feature    134137. .151328
                     /note="assembly_fragment:01324
                     fragment_chain:3"
     misc_feature    151429. .167242
                     /note="assembly_fragment:01775
                     fragment_chain:3"
     misc_feature    167343. .179999
                     /note="assembly_fragment:00411
                     fragment_chain:3"
     misc_feature    180100. .190992
                     /note="assembly_fragment:00565.0"
     misc_feature    191093. .228676
                     /note="assembly_fragment:02360"
ORIGIN

  Query Match             2.3%;  Score 26;  DB 2;  Length 228676;
  Best Local Similarity   100.0%;  Pred. No. 0.018;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        867 AAAAAAAGAAAAAAAAACTGTTCAAT 892
              ||||||||||||||||||||||||||
Db     122236 AAAAAAAGAAAAAAAAACTGTTCAAT 122211


RESULT 8
AC021823/c
LOCUS       AC021823              146015 bp    DNA     linear   HTG 04-APR-2000
DEFINITION  Homo sapiens chromosome 2 clone RP11-169P13 map 2, WORKING DRAFT
            SEQUENCE, 22 unordered pieces.
ACCESSION   AC021823
VERSION     AC021823.3  GI:7408019
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 146015)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 2, clone RP11-169P13
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 146015)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JAN-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Apr 4, 2000 this sequence version replaced gi:6957782.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L5939
                Center clone name: 169_P_13
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 135544 bases at least Q40
                Consensus quality: 140389 bases at least Q30
                Consensus quality: 142539 bases at least Q20
                Insert size: 153000; agarose-fp
                Insert size: 143915; sum-of-contigs
                Quality coverage: 4.1 in Q20 bases; agarose-fp
                Quality coverage: 4.4 in Q20 bases; sum-of-contigs
```

```
            ----------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 22 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1       40: contig of 40 bp in length
            *       41      140: gap of 100 bp
            *      141     1665: contig of 1525 bp in length
            *     1666     1765: gap of 100 bp
            *     1766     3194: contig of 1429 bp in length
            *     3195     3294: gap of 100 bp
            *     3295     5598: contig of 2304 bp in length
            *     5599     5698: gap of 100 bp
            *     5699    10030: contig of 4332 bp in length
            *    10031    10130: gap of 100 bp
            *    10131    13650: contig of 3520 bp in length
            *    13651    13750: gap of 100 bp
            *    13751    18728: contig of 4978 bp in length
            *    18729    18828: gap of 100 bp
            *    18829    21760: contig of 2932 bp in length
            *    21761    21860: gap of 100 bp
            *    21861    25279: contig of 3419 bp in length
            *    25280    25379: gap of 100 bp
            *    25380    29185: contig of 3806 bp in length
            *    29186    29285: gap of 100 bp
            *    29286    33428: contig of 4143 bp in length
            *    33429    33528: gap of 100 bp
            *    33529    39961: contig of 6433 bp in length
            *    39962    40061: gap of 100 bp
            *    40062    43365: contig of 3304 bp in length
            *    43366    43465: gap of 100 bp
            *    43466    48258: contig of 4793 bp in length
            *    48259    48358: gap of 100 bp
            *    48359    54546: contig of 6188 bp in length
            *    54547    54646: gap of 100 bp
            *    54647    65171: contig of 10525 bp in length
            *    65172    65271: gap of 100 bp
            *    65272    78708: contig of 13437 bp in length
            *    78709    78808: gap of 100 bp
            *    78809    90249: contig of 11441 bp in length
            *    90250    90349: gap of 100 bp
            *    90350   102852: contig of 12503 bp in length
            *   102853   102952: gap of 100 bp
            *   102953   114466: contig of 11514 bp in length
            *   114467   114566: gap of 100 bp
            *   114567   127728: contig of 13162 bp in length
            *   127729   127828: gap of 100 bp
            *   127829   146015: contig of 18187 bp in length.
FEATURES             Location/Qualifiers
     source          1. .146015
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2"
                     /clone="RP11-169P13"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .40
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     misc_feature    141. .1665
                     /note="assembly_fragment"
     misc_feature    1766. .3194
                     /note="assembly_fragment"
     misc_feature    3295. .5598
                     /note="assembly_fragment"
     misc_feature    5699. .10030
                     /note="assembly_fragment"
     misc_feature    10131. .13650
                     /note="assembly_fragment"
     misc_feature    13751. .18728
                     /note="assembly_fragment"
     misc_feature    18829. .21760
                     /note="assembly_fragment"
     misc_feature    21861. .25279
                     /note="assembly_fragment"
     misc_feature    25380. .29185
                     /note="assembly_fragment"
     misc_feature    29286. .33428
                     /note="assembly_fragment"
     misc_feature    33529. .39961
                     /note="assembly_fragment"
     misc_feature    40062. .43365
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:right"
     misc_feature    43466. .48258
                     /note="assembly_fragment"
     misc_feature    48359. .54546
                     /note="assembly_fragment"
     misc_feature    54647. .65171
                     /note="assembly_fragment"
     misc_feature    65272. .78708
                     /note="assembly_fragment"
     misc_feature    78809. .90249
                     /note="assembly_fragment"
     misc_feature    90350. .102852
                     /note="assembly_fragment"
     misc_feature    102953. .114466
                     /note="assembly_fragment"
     misc_feature    114567. .127728
                     /note="assembly_fragment"
     misc_feature    127829. .146015
                     /note="assembly_fragment"
ORIGIN

  Query Match             2.2%;  Score 24;  DB 2;  Length 146015;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        865 CTAAAAAAAGAAAAAAAAACTGTT 888
              ||||||||||||||||||||||||
Db     112269 CTAAAAAAAGAAAAAAAAACTGTT 112246


RESULT 9
AC092642
LOCUS       AC092642              152251 bp    DNA     linear   PRI 01-MAR-2002
DEFINITION  Homo sapiens BAC clone RP11-391P1 from 2, complete sequence.
ACCESSION   AC092642 AC023761
VERSION     AC092642.2  GI:15638834
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 152251)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
   PUBMED   9847074
REFERENCE   2  (bases 1 to 152251)
  AUTHORS   Mulvaney,E., Abbott,A., Dixon,R., Dignan,G. and Phillips,A.
  TITLE     The sequence of Homo sapiens BAC clone RP11-391P1
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 152251)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
```

```
            MO 63108, USA
REFERENCE   4  (bases 1 to 152251)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-SEP-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 152251)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAR-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Sep 18, 2001 this sequence version replaced gi:14916226.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0391P01
                Drafting Center: WIBR
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. John D.
            McPherson, Department of Genetics, Washington University, St. Louis
            MO. For additional information about the map position of this
            sequence, see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries. Genomics 51:1-8. The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at http://www.chori.org
            VECTOR:  pBACe3.6

            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is RP11-544E11, 2000 bp overlap;
            the clone sequenced to the right is RP11-710J17, 2000 bp overlap.
            Actual end of this clone is at base position 19440 of RP11-710J17.

            Data from AC062033, AC061960, and AC021823 was used to finish this
            clone, AC023761.  Polymorphisms have been identified between
            AC062033, AC061960, and AC023761.

            The sequence of AC023761 has been incorporated into AC092642.
FEATURES             Location/Qualifiers
     source          1. .152251
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2"
                     /clone="RP11-391P1"
                     /clone_lib="RPCI-11"
     repeat_region   4. .169
                     /rpt_family="MaLR"
     repeat_region   206. .495
                     /rpt_family="MaLR"
     repeat_region   563. .718
                     /rpt_family="L2"
     repeat_region   1151. .1293
                     /rpt_family="L2"
     repeat_region   1524. .1624
                     /rpt_family="L1"
     repeat_region   1622. .1643
                     /rpt_family="AT_rich"
     repeat_region   1628. .1694
                     /rpt_family="Alu"
     repeat_region   1695. .2139
                     /rpt_family="L1"
     repeat_region   1794. .1837
                     /rpt_family="(TG)n"
     repeat_region   2149. .2433
                     /rpt_family="MaLR"
     repeat_region   2449. .3161
                     /rpt_family="L1"
     repeat_region   3162. .3456
                     /rpt_family="Alu"
     repeat_region   3325. .3365
                     /rpt_family="(T)n"
     repeat_region   3457. .3758
                     /rpt_family="L1"
     repeat_region   3759. .4066
                     /rpt_family="Alu"
     repeat_region   4067. .4492
                     /rpt_family="L1"
     repeat_region   4493. .4779
                     /rpt_family="Alu"
     repeat_region   4780. .6269
                     /rpt_family="L1"
     repeat_region   6300. .6363
                     /rpt_family="L1"
     repeat_region   6367. .6448
                     /rpt_family="ERVL"
     repeat_region   6470. .6731
                     /rpt_family="ERVL"
     repeat_region   6737. .6876
                     /rpt_family="L1"
     repeat_region   7608. .7643
                     /rpt_family="(CA)n"
     repeat_region   7956. .8394
                     /rpt_family="MaLR"
     repeat_region   9284. .9382
                     /rpt_family="L1"
     repeat_region   9480. .9537
                     /rpt_family="AT_rich"
     repeat_region   9534. .9825
                     /rpt_family="Alu"
     misc_feature    9934
                     /note="match to EST AI038041 (NID:g3277235) ox29h03.x1"
     misc_feature    10044. .10443
                     /note="match to EST AI038041 (NID:g3277235) ox29h03.x1"
     repeat_region   10054. .10075
                     /rpt_family="AT_rich"
     misc_feature    10780. .11226
                     /note="match to EST T91378 (NID:g723291) yd53a07.s1"
     misc_feature    10799. .10943
                     /note="match to EST H81620 (NID:g1059709) ys66d09.s1"
     repeat_region   10853. .10901
                     /rpt_family="(TG)n"
     misc_feature    11080. .11325
                     /note="match to EST H81620 (NID:g1059709) ys66d09.s1"
     misc_feature    11246. .11379
                     /note="match to EST T84526 (NID:g712878) yd53a07.r1"
     repeat_region   11828. .11945
                     /rpt_family="L2"
     misc_feature    12091. .12249
                     /note="match to EST H81619 (NID:g1059708) ys66d09.r1"
     misc_feature    12098. .12249
```

```
                     /note="match to EST T84526 (NID:g712878) yd53a07.r1"
     repeat_region   12315. .12336
                     /rpt_family="AT_rich"
     repeat_region   15388. .15522
                     /rpt_family="L2"
     repeat_region   15624. .15788
                     /rpt_family="(TG)n"
     repeat_region   15831. .15898
                     /rpt_family="(TG)n"
     repeat_region   15934. .16232
                     /rpt_family="Alu"
     repeat_region   18502. .18571
                     /rpt_family="L2"
     repeat_region   19347. .19607
                     /rpt_family="Alu"
     repeat_region   20299. .20328
                     /rpt_family="AT_rich"
     repeat_region   20300. .20413
                     /rpt_family="L1"
     repeat_region   20821. .21353
                     /rpt_family="L2"
     repeat_region   21356. .21468
                     /rpt_family="L1"
     repeat_region   21449. .21505
                     /rpt_family="AT_rich"
     repeat_region   21504. .21678
                     /rpt_family="MER1_type"
     repeat_region   22507. .22594
                     /rpt_family="AcHobo"
     repeat_region   22644. .23228

  Query Match             2.2%;  Score 24;  DB 9;  Length 152251;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        865 CTAAAAAAAGAAAAAAAAACTGTT 888
              ||||||||||||||||||||||||
Db      26051 CTAAAAAAAGAAAAAAAAACTGTT 26074


RESULT 10
BX927299
LOCUS       BX927299              155054 bp    DNA     linear   HTG 08-OCT-2004
DEFINITION  Danio rerio clone DKEY-146H10, WORKING DRAFT SEQUENCE, 7 unordered
            pieces.
ACCESSION   BX927299
VERSION     BX927299.5  GI:54019870
KEYWORDS    HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 155054)
  AUTHORS   Heath,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-OCT-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Oct 9, 2004 this sequence version replaced gi:52313301.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -------------- Project Information
            Center project name: zK146H10
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 153838 bases at least Q40
            Consensus quality: 153985 bases at least Q30
            Consensus quality: 154096 bases at least Q20
            Insert size: 154454; sum-of-contigs
            Insert size: 158561; 1.9% error; agarose-fp
            Quality coverage: 11.18x in Q20 bases; sum-of-contigs Quality
            coverage: 10.90x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 7 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    33656: contig of 33656 bp in length
            *    33657    33756: gap of 100 bp
            *    33757    51099: contig of 17343 bp in length
            *    51100    51199: gap of 100 bp
            *    51200    60096: contig of 8897 bp in length
            *    60097    60196: gap of 100 bp
            *    60197    85868: contig of 25672 bp in length
            *    85869    85968: gap of 100 bp
            *    85969   116591: contig of 30623 bp in length
            *   116592   116691: gap of 100 bp
            *   116692   147668: contig of 30977 bp in length
            *   147669   147768: gap of 100 bp
            *   147769   155054: contig of 7286 bp in length.
FEATURES             Location/Qualifiers
     source          1. .155054
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-146H10"
                     /clone_lib="DanioKey"
     misc_feature    1. .33656
                     /note="assembly_fragment:00001
                     fragment_chain:1"
     misc_feature    33757. .51099
                     /note="assembly_fragment:02384
                     fragment_chain:1"
     misc_feature    51200. .60096
                     /note="assembly_fragment:02382
                     fragment_chain:2"
     misc_feature    60197. .85868
                     /note="assembly_fragment:01606
                     fragment_chain:2"
     misc_feature    85969. .116591
                     /note="assembly_fragment:00526.0"
     misc_feature    116692. .147668
                     /note="assembly_fragment:01071"
     misc_feature    147769. .155054
                     /note="assembly_fragment:02238"
ORIGIN

  Query Match             2.2%;  Score 24;  DB 2;  Length 155054;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        868 AAAAAAGAAAAAAAAACTGTTCAA 891
              ||||||||||||||||||||||||
Db      92426 AAAAAAGAAAAAAAAACTGTTCAA 92449


RESULT 11
BX001022/c
LOCUS       BX001022              160804 bp    DNA     linear   VRT 11-OCT-2003
DEFINITION  Zebrafish DNA sequence from clone CH211-243O19, complete sequence.
ACCESSION   BX001022
VERSION     BX001022.9  GI:37652295
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
```

```
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 160804)
  AUTHORS   Giselle,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-OCT-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Oct 13, 2003 this sequence version replaced gi:35209032.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU).  For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml
            CH211-243O19 is from a CHORI-211 BAC library
            VECTOR: pTARBAC2.1.
FEATURES             Location/Qualifiers
     source          1. .160804
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="CH211-243O19"
                     /clone_lib="CHORI-211"

ORIGIN

  Query Match             2.2%;  Score 24;  DB 5;  Length 160804;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        868 AAAAAAGAAAAAAAAACTGTTCAA 891
              ||||||||||||||||||||||||
Db      44746 AAAAAAGAAAAAAAAACTGTTCAA 44723


RESULT 12
BX927400/c
LOCUS       BX927400              185217 bp    DNA     linear   HTG 10-OCT-2004
DEFINITION  Danio rerio clone CH211-195K18, WORKING DRAFT SEQUENCE.
ACCESSION   BX927400
VERSION     BX927400.6  GI:51965253
KEYWORDS    HTG; HTGS_PHASE2; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 185217)
  AUTHORS   Ellwood,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-OCT-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Sep 9, 2004 this sequence version replaced gi:51592026.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -------------- Project Information
            Center project name: zC195K18
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 184890 bases at least Q40
            Consensus quality: 185061 bases at least Q30
            Consensus quality: 185152 bases at least Q20
            Insert size: 185217; sum-of-contigs
            Insert size: 210440; 9.8% error; agarose-fp
            Quality coverage: 9.78x in Q20 bases; sum-of-contigs Quality
            coverage: 8.61x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 1 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *       1   185217: contig of 185217 bp in length.
FEATURES             Location/Qualifiers
     source          1. .185217
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="CH211-195K18"
                     /clone_lib="CHORI-211"
     misc_feature    1. .185217
                     /note="assembly_fragment:02923
                     clone_end:SP6
                     vector_side:right
                     clone_end:T7
                     vector_side:left"
ORIGIN

  Query Match             2.2%;  Score 24;  DB 2;  Length 185217;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        868 AAAAAAGAAAAAAAAACTGTTCAA 891
              ||||||||||||||||||||||||
Db       8803 AAAAAAGAAAAAAAAACTGTTCAA 8780


RESULT 13
BX649502/c
LOCUS       BX649502              247387 bp    DNA     linear   VRT 06-DEC-2003
DEFINITION  Zebrafish DNA sequence from clone DKEY-204F11 in linkage group 3,
            complete sequence.
ACCESSION   BX649502
VERSION     BX649502.4  GI:39540484
KEYWORDS    HTG.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
```

```
REFERENCE   1  (bases 1 to 247387)
  AUTHORS   Whitehead,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-DEC-2003) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Dec 6, 2003 this sequence version replaced gi:38201304.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            --------------
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest, except on the rare
            occasion of the clone being a YAC.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived
            Zebrafish pUC subclones occasionally display inconsistency over the
            length of mononucleotide A/T runs and conserved TA repeats. Where
            this is found the longest good quality representation will be
            submitted.
            Repeat names beginning 'Dr' were identified by the Recon repeat
            discovery system (Zhirong Bao and Sean Eddy, submitted), and those
            beginning 'drr' were identified by Rick Waterman (Stephen Johnson
            lab, WashU).  For further information see
            http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEY-204F11
            is from a Zebrafish BAC library
            VECTOR: pIndigoBAC-5.
FEATURES             Location/Qualifiers
     source          1. .247387
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-204F11"
                     /clone_lib="DanioKey"
ORIGIN

  Query Match             2.2%;  Score 24;  DB 5;  Length 247387;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        867 AAAAAAAGAAAAAAAAACTGTTCA 890
              ||||||||||||||||||||||||
Db      64795 AAAAAAAGAAAAAAAAACTGTTCA 64772


RESULT 14
BX927111/c
LOCUS       BX927111              263169 bp    DNA     linear   HTG 05-APR-2004
DEFINITION  Danio rerio clone DKEY-54K13, *** SEQUENCING IN PROGRESS ***, 12
            unordered pieces.
ACCESSION   BX927111
VERSION     BX927111.5  GI:46237716
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 263169)
  AUTHORS   Sims,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-APR-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Apr 6, 2004 this sequence version replaced gi:46200415.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -------------- Project Information
            Center project name: zK54K13
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 260222 bases at least Q40
            Consensus quality: 260911 bases at least Q30
            Consensus quality: 261370 bases at least Q20
            Insert size: 262069; sum-of-contigs
            Insert size: 249466; 4.6% error; agarose-fp
            Quality coverage: 8.81x in Q20 bases; sum-of-contigs Quality
            coverage: 9.36x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 12 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     8720: contig of 8720 bp in length
            *     8721     8820: gap of 100 bp
            *     8821    34199: contig of 25379 bp in length
            *    34200    34299: gap of 100 bp
            *    34300    61537: contig of 27238 bp in length
            *    61538    61637: gap of 100 bp
            *    61638    82341: contig of 20704 bp in length
            *    82342    82441: gap of 100 bp
            *    82442   113587: contig of 31146 bp in length
            *   113588   113687: gap of 100 bp
            *   113688   117389: contig of 3702 bp in length
            *   117390   117489: gap of 100 bp
            *   117490   169953: contig of 52464 bp in length
            *   169954   170053: gap of 100 bp
            *   170054   191893: contig of 21840 bp in length
            *   191894   191993: gap of 100 bp
            *   191994   232102: contig of 40109 bp in length
            *   232103   232202: gap of 100 bp
            *   232203   235712: contig of 3510 bp in length
            *   235713   235812: gap of 100 bp
            *   235813   256314: contig of 20502 bp in length
            *   256315   256414: gap of 100 bp
            *   256415   263169: contig of 6755 bp in length.
FEATURES             Location/Qualifiers
     source          1. .263169
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-54K13"
                     /clone_lib="DanioKey"
     misc_feature    1. .8720
                     /note="assembly_fragment:00210
                     fragment_chain:1"
     misc_feature    8821. .34199
                     /note="assembly_fragment:01606
                     fragment_chain:1"
     misc_feature    34300. .61537
                     /note="assembly_fragment:01245
                     fragment_chain:1"
```

```
     misc_feature    61638. .82341
                     /note="assembly_fragment:00629
                     fragment_chain:1"
     misc_feature    82442. .113587
                     /note="assembly_fragment:01994
                     fragment_chain:1"
     misc_feature    113688. .117389
                     /note="assembly_fragment:00023
                     fragment_chain:1"
     misc_feature    117490. .169953
                     /note="assembly_fragment:03001
                     fragment_chain:1"
     misc_feature    170054. .191893
                     /note="assembly_fragment:00915
                     fragment_chain:1"
     misc_feature    191994. .232102
                     /note="assembly_fragment:02450
                     fragment_chain:1"
     misc_feature    232203. .235712
                     /note="assembly_fragment:00058
                     fragment_chain:1"
     misc_feature    235813. .256314
                     /note="assembly_fragment:00359
                     fragment_chain:1"
     misc_feature    256415. .263169
                     /note="assembly_fragment:00104
                     fragment_chain:1"
ORIGIN

  Query Match             2.2%;  Score 24;  DB 2;  Length 263169;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        867 AAAAAAAGAAAAAAAAACTGTTCA 890
              ||||||||||||||||||||||||
Db     251704 AAAAAAAGAAAAAAAAACTGTTCA 251681


RESULT 15
AC128496/c
LOCUS       AC128496               291544 bp    DNA     linear   HTG 19-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-444D9, WORKING DRAFT SEQUENCE, 3
            unordered pieces.
ACCESSION   AC128496
VERSION     AC128496.3  GI:25085139
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 291544)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 291544)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 291544)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 19, 2002 this sequence version replaced gi:23907748.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GYSW
                Center clone name: CH230-444D9
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 187201 bases at least Q40
                Consensus quality: 189849 bases at least Q30
                Consensus quality: 191486 bases at least Q20
                Estimated insert size: 193264; sum-of-contigs estimation
                Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
```

```
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1   288506: contig of 288506 bp in length
            *  288507   288606: gap of unknown length
            *  288607   289949: contig of 1343 bp in length
            *  289950   290049: gap of unknown length
            *  290050   291544: contig of 1495 bp in length.
FEATURES             Location/Qualifiers
     source          1. .291544
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-444D9"
     misc_feature    1. .1841
                     /note="wgs_end_extension
                     clone_end:T7"
     misc_feature    6777. .7677
                     /note="clone_boundary
                     clone_end:T7
                     site:
                     end_sequence:BZ198832"
     misc_feature    complement(187560. .188645)
                     /note="clone_boundary
                     clone_end:Sp6
                     site:
                     end_sequence:BZ198833"
     misc_feature    189481. .192835
                     /note="wgs_end_extension
                     clone_end:Sp6"
     misc_feature    195009. .196648
                     /note="wgs_end_extension
                     clone_end:Sp6"
ORIGIN

  Query Match             2.2%;  Score 24;  DB 2;  Length 291544;
  Best Local Similarity   100.0%;  Pred. No. 0.24;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        867 AAAAAAAGAAAAAAAAACTGTTCA 890
              ||||||||||||||||||||||||
Db     130151 AAAAAAAGAAAAAAAAACTGTTCA 130128


Search completed: August 29, 2005, 04:52:49
Job time : 5169 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 29, 2005, 00:19:31 ; Search time 701 Seconds
(without alignments)
9382.072 Million cell updates/sec

Title: US-09-979-549-2
Perfect score: 1111
Sequence: 1 gaagtcagaaggccgttcag.....ccgagaagagagagagagag 1111

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1111	100.0	2687	5	AAC84132	Aac84132 Wheat TaT
	2	22	2.0	141	6	ABL86452	Abl86452 Human ova
	3	22	2.0	144	6	ABL86187	Abl86187 Human ova
	4	22	2.0	154	6	ABL86321	Abl86321 Human ova
	5	22	2.0	213	6	ABL86948	Abl86948 Human ova
c	6	22	2.0	1074	5	ADL62118	Adl62118 Human ova
c	7	22	2.0	24935	4	ABL10390	Abl10390 Drosophil
c	8	21	1.9	377	2	AAV86132	Aav86132 EST clone
	9	21	1.9	403	3	AAC24900	Aac24900 Human sec
	10	21	1.9	621	4	AAH71551	Aah71551 Human cer
c	11	21	1.9	1097	3	AAF16195	Aaf16195 Human pro
c	12	21	1.9	2000	8	ADA72818	Ada72818 Rice gene
c	13	21	1.9	2170	4	AAH72880	Aah72880 Human cer
c	14	21	1.9	2170	5	ABV22589	Abv22589 Human pro
c	15	21	1.9	2170	5	ABV25695	Abv25695 Human pro
c	16	21	1.9	2170	5	ABV28409	Abv28409 Human pro
c	17	21	1.9	2170	5	ABV24749	Abv24749 Human pro
c	18	21	1.9	3191	3	AAA16647	Aaa16647 Human sec
c	19	21	1.9	3618	6	ABL56484	Abl56484 Nucleotid
c	20	21	1.9	4522	3	AAC76440	Aac76440 Human ORF
c	21	21	1.9	4740	5	ABV21724	Abv21724 Human pro
c	22	21	1.9	4740	5	ABV27546	Abv27546 Human pro
c	23	21	1.9	10351	6	ABK88932	Abk88932 Human CD1
c	24	21	1.9	80815	13	ABD33381	Abd33381 Human can
	25	21	1.9	254396	12	ADQ97557	Adq97557 Human can
	26	20	1.8	162	6	ABZ08670	Abz08670 Human leu
c	27	20	1.8	371	6	ABK64451	Abk64451 Human ben
	28	20	1.8	375	5	ABV48688	Abv48688 Human pro
	29	20	1.8	419	4	AAI85788	Aai85788 Human pol
	30	20	1.8	422	6	ABQ58941	Abq58941 Human col
c	31	20	1.8	461	2	AAV89657	Aav89657 EST clone
	32	20	1.8	462	9	ACH22705	Ach22705 Human adu
	33	20	1.8	555	4	AAS36746	Aas36746 Human car
	34	20	1.8	555	10	ADE47440	Ade47440 Human car
	35	20	1.8	555	13	ADJ08858	Adj08858 Human car
c	36	20	1.8	603	4	AAK52328	Aak52328 Human pol
c	37	20	1.8	722	6	ADB87555	Adb87555 Plasmid p
c	38	20	1.8	860	4	AAH03455	Aah03455 Human cDN
c	39	20	1.8	870	3	AAA08590	Aaa08590 Human cyt
c	40	20	1.8	1952	12	ADI42697	Adi42697 Plant tra
c	41	20	1.8	1952	12	ADO02953	Ado02953 Soybean o
	42	20	1.8	2143	11	ACN92902	Acn92902 Breast ca
	43	20	1.8	2559	13	ACN38938	Acn38938 Tumour-as
	44	20	1.8	2577	2	AAX87624	Aax87624 Set* spli
	45	20	1.8	2577	6	ABS73234	Abs73234 DNA encod

ALIGNMENTS

RESULT 1
AAC84132
ID AAC84132 standard; DNA; 2687 BP.
XX
AC AAC84132;
XX
DT 09-APR-2001 (first entry)
XX
DE Wheat TaTrxh2 gene.
XX
KW Wheat; TaTrxh2; thioredoxin; promoter; gene expression; transgenic plant;
KW monocotyledon; plant cell; seed; amylaceous albumen; ds.
XX
OS Triticum aestivum.
XX

FH	Key	Location/Qualifiers
FT	promoter	1. .1111
FT		/*tag= a
FT		/note= "promoter region is specifically claimed"
FT	misc_signal	550. .558
FT		/*tag= b
FT		/label= Gibberellic_acid_response_element
FT		/function= "regulates gene expression in response to gibberellic acid"
FT	misc_signal	561. .569
FT		/*tag= c
FT		/label= Gibberellic_acid_response_element
FT		/function= "regulates gene expression in response to gibberellic acid"
FT	protein_bind	699. .708
FT		/*tag= d
FT		/bound_moiety= "GCN4-like protein"
FT		/label= GCNA-like_box
FT	protein_bind	860
FT		/*tag= e
FT		/bound_moiety= "leucine zipper proteins"
FT		/label= bzip_motif
FT	misc_signal	867. .883
FT		/*tag= f
FT		/function= "possible regulator of gene expression in response to abscisic acid"
FT		/note= "present in thioredoxin h gene promoter sequences from tobacco and rice"

```
FT   protein_bind    901. .906
FT                   /*tag= g
FT                   /bound_moiety= "basic helix-loop-helix transcription
FT                   factor"
FT                   /label= bHLH_recognition_motif
FT   protein_bind    924. .927
FT                   /*tag= h
FT                   /bound_moiety= "leucine zipper proteins"
FT                   /label= bzip_motif
FT   misc_signal     1001. .1003
FT                   /*tag= i
FT                   /function= "involved in aleurone layer-specific gene
FT                   expression"
FT   misc_signal     1002. .1008
FT                   /*tag= j
FT                   /note= "TATA-like box"
FT   misc_signal     1011. .1013
FT                   /*tag= k
FT                   /function= "involved in aleurone layer-specific gene
FT                   expression"
FT   protein_bind    1021. .1028
FT                   /*tag= l
FT                   /bound_moiety= "Sp1 transcription factor"
FT                   /label= GC_box
FT   exon            1047. .1231
FT                   /*tag= m
FT                   /number= 1
FT   CDS             1112. .2557
FT                   /*tag= n
FT                   /gene= "TaTrxh2"
FT                   /product= "thioredoxin h2"
FT                   /note= "CDS contains introns"
FT   intron          1232. .2202
FT                   /*tag= o
FT                   /number= 1
FT   exon            2203. .2325
FT                   /*tag= p
FT                   /number= 2
FT   intron          2326. .2428
FT                   /*tag= q
FT                   /number= 2
FT                   /cons_splice= (5'site:YES, 3'site:NO)
FT   exon            2429. .2687
FT                   /*tag= r
FT                   /number= 3
FT   3'UTR           2558. .2687
FT                   /*tag= s
FT   polyA_signal    2584. .2589
FT                   /*tag= t
XX
PN   WO200070065-A1.
XX
PD   23-NOV-2000.
XX
PF   17-MAY-2000; 2000WO-FR001318.
XX
PR   17-MAY-1999;   99FR-00006231.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX
PI   Gautier M,  Ihorai T,  Joudrier P;
XX
DR   WPI; 2001-016241/02.
DR   P-PSDB; AAB37102.
XX
PT   New promoter from a wheat thioredoxin gene, useful for controlling
PT   transgene expression in plants, provides seed-specific expression.
XX
PS   Example 2; Page 28-30; 34pp; French.
XX
CC   This sequence corresponds to the wheat (Triticum aestivum) TaTrxh2 which
CC   encodes a thioredoxin h protein of 126 amino acids, molecular mass 13435
CC   daltons and an isoelectric focussing point (pI) 5.0. The promoter is used
CC   to control gene expression in transgenic plants (particularly
CC   monocotyledons) or plant cells, especially for seed-specific expression,
CC   particularly in the amylaceous albumen
XX
SQ   Sequence 2687 BP; 638 A; 626 C; 690 G; 733 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 1111;  DB 5;  Length 2687;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1111;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GAAGTCAGAAGGCCGTTCAGAATTGTTGGAGGACTCGAAAAAAAGAAGGGGAGCCCAGGC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GAAGTCAGAAGGCCGTTCAGAATTGTTGGAGGACTCGAAAAAAAGAAGGGGAGCCCAGGC 60

Qy         61 AGACGACGGGGCGGCATGTGCCTGTTCCTTGGCGAGGCGTCTAGCTTTGGCAGCCGCCGC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGACGACGGGGCGGCATGTGCCTGTTCCTTGGCGAGGCGTCTAGCTTTGGCAGCCGCCGC 120

Qy        121 CGCTTTTCTCCTTGGGTGGGCGCGCGAGCTCCCCGAGTTTGAGCCGCAATTTTTTTACAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CGCTTTTCTCCTTGGGTGGGCGCGCGAGCTCCCCGAGTTTGAGCCGCAATTTTTTTACAT 180

Qy        181 TTTATGGCGATGGCGTCAGGCGTTTATCTAGGCGTCTGGGAGGGTACATTTGAAGATGTG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTTATGGCGATGGCGTCAGGCGTTTATCTAGGCGTCTGGGAGGGTACATTTGAAGATGTG 240

Qy        241 CCACCAACTCCAAACCGACAACCCTGTATCTGAGCATGCCTCATGCCTCTCCTTCATGCC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCACCAACTCCAAACCGACAACCCTGTATCTGAGCATGCCTCATGCCTCTCCTTCATGCC 300

Qy        301 TCCCTTTGGGTGAGGTCATGTGCCCTTGGCGGCGAGTGGCTTCCCGTTTAGAGCAAGTAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCCCTTTGGGTGAGGTCATGTGCCCTTGGCGGCGAGTGGCTTCCCGTTTAGAGCAAGTAT 360

Qy        361 AATAAGTCCTAGTCAGCTGGCTATAAGATGTTCCACATCAGCAAATCCTTAAACTGGAGG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AATAAGTCCTAGTCAGCTGGCTATAAGATGTTCCACATCAGCAAATCCTTAAACTGGAGG 420

Qy        421 AGAAAGAAAGTAGGAGTGAGAAGGGCGTCGGCGCTTCGTCAATCGCTAGCGATAGCACAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AGAAAGAAAGTAGGAGTGAGAAGGGCGTCGGCGCTTCGTCAATCGCTAGCGATAGCACAA 480

Qy        481 GCTCCCATGGAATCGAGCCAACATGCAACCCGCACAATGACTAAAGGCAAACGCCAGCCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCTCCCATGGAATCGAGCCAACATGCAACCCGCACAATGACTAAAGGCAAACGCCAGCCA 540

Qy        541 ATCAGTATGCCTTTCTCTGCATCTTTCTTCATGCAAGCATTAAATACTATAGCTAATCTA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATCAGTATGCCTTTCTCTGCATCTTTCTTCATGCAAGCATTAAATACTATAGCTAATCTA 600

Qy        601 CAGCCAGTTTATTATATAAACAGGCTATATAGCTGACCTGGCAGTGCTATAGAGCCGGCA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CAGCCAGTTTATTATATAAACAGGCTATATAGCTGACCTGGCAGTGCTATAGAGCCGGCA 660

Qy        661 GCCGGCTCTTCTATTAGCTTTGCTCTTATGGCTACATCTGTGTGAGCAGTCGATTGATTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCGGCTCTTCTATTAGCTTTGCTCTTATGGCTACATCTGTGTGAGCAGTCGATTGATTC 720

Qy        721 AAACAACAAATCCGGGCGTTCAGCAAGTCGGAATGAATTTCGGCTCATCACTCATTGTCG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAACAACAAATCCGGGCGTTCAGCAAGTCGGAATGAATTTCGGCTCATCACTCATTGTCG 780

Qy        781 TGGGCCTCACGCGTATTCGCCTAACCGTGTTTGAATCAGACCCTCACGAAGCCACGGCTC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGGGCCTCACGCGTATTCGCCTAACCGTGTTTGAATCAGACCCTCACGAAGCCACGGCTC 840

Qy        841 CAGCGACCCGTTCACCACGTCAGCCTAAAAAAAGAAAAAAAAACTGTTCAATCACACGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CAGCGACCCGTTCACCACGTCAGCCTAAAAAAAGAAAAAAAAACTGTTCAATCACACGCC 900

Qy        901 CATCTGAACCGTTCAACAGCCCCACGTAATTTCGCGCACCAGCAAAGGGCATATCCGTCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 CATCTGAACCGTTCAACAGCCCCACGTAATTTCGCGCACCAGCAAAGGGCATATCCGTCA 960
```

```
Qy        961 TAGCGAGCGCATAAATTCTGATTCCTGCCTGCCTGCCGGACAATTTATCTTTGGGGAGGC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TAGCGAGCGCATAAATTCTGATTCCTGCCTGCCTGCCGGACAATTTATCTTTGGGGAGGC 1020

Qy       1021 GGGCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAATCAACA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGGCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAATCAACA 1080

Qy       1081 AGCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 1111
              ||||||||||||||||||||||||||||||||
Db       1081 AGCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 1111
```

RESULT 2
ABL86452
ID ABL86452 standard; cDNA; 141 BP.
XX
AC ABL86452;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human ovarian cancer related cDNA clone SEQ ID NO:9430.
XX
KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200192581-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US017756.
XX
PR 26-MAY-2000; 2000US-0207484P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Harlocker SL, Jones R;
XX
DR WPI; 2002-122075/16.
XX
PT Composition for therapy and diagnosis of ovarian cancer comprising
PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT polypeptide, antibody specific to polypeptide or T cell expressing
PT polypeptide.
XX
PS Claim 1; SEQ ID NO 9430; 489pp; English.
XX
CC The present invention describes a composition (I) comprising: carriers
CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
CC from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
CC (III) encoding (II) having a sequence (S2), a T cell population of (II),
CC or antigen presenting cells that express (II). (I) has cytostatic
CC activity. An oligonucleotide (IV) that hybridises to (S1) can be used for
CC detecting ovarian cancer in a patient's biological sample preferably
CC serum or ovarian tissue. The method comprises contacting a biological
CC sample from a patient with (IV), detecting the amount of polynucleotide
CC hybridising to (IV) and comparing the amount to a predetermined cutoff
CC value and thereby detecting ovarian cancer in the patient, where the
CC amount of polynucleotide hybridising to (IV) is detected preferably by
CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
CC useful for stimulating and/or expanding T cells specific for an ovarian
CC tumour protein comprising contacting T cells with (III) or (II). (III) is
CC useful in design and preparation of ribozyme molecules for inhibiting
CC expression of the tumour polypeptides and proteins in tumour cells; and
CC to isolate a full length gene from a suitable library e.g., a tumour cDNA
CC library using well known techniques
XX
SQ Sequence 141 BP; 58 A; 26 C; 46 G; 11 T; 0 U; 0 Other;

Query Match 2.0%; Score 22; DB 6; Length 141;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        862 AGCCTAAAAAAAGAAAAAAAAA 883
              |||||||||||||||||||||||
Db         98 AGCCTAAAAAAAGAAAAAAAAA 119
```

RESULT 3
ABL86187
ID ABL86187 standard; cDNA; 144 BP.
XX
AC ABL86187;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human ovarian cancer related cDNA clone SEQ ID NO:9165.
XX
KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200192581-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US017756.
XX
PR 26-MAY-2000; 2000US-0207484P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Harlocker SL, Jones R;
XX
DR WPI; 2002-122075/16.
XX
PT Composition for therapy and diagnosis of ovarian cancer comprising
PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT polypeptide, antibody specific to polypeptide or T cell expressing
PT polypeptide.
XX
PS Claim 1; SEQ ID NO 9165; 489pp; English.
XX
CC The present invention describes a composition (I) comprising: carriers
CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
CC from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
CC (III) encoding (II) having a sequence (S2), a T cell population of (II),
CC or antigen presenting cells that express (II). (I) has cytostatic
CC activity. An oligonucleotide (IV) that hybridises to (S1) can be used for
CC detecting ovarian cancer in a patient's biological sample preferably
CC serum or ovarian tissue. The method comprises contacting a biological
CC sample from a patient with (IV), detecting the amount of polynucleotide
CC hybridising to (IV) and comparing the amount to a predetermined cutoff
CC value and thereby detecting ovarian cancer in the patient, where the
CC amount of polynucleotide hybridising to (IV) is detected preferably by
CC polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
CC useful for stimulating and/or expanding T cells specific for an ovarian
CC tumour protein comprising contacting T cells with (III) or (II). (III) is
CC useful in design and preparation of ribozyme molecules for inhibiting
CC expression of the tumour polypeptides and proteins in tumour cells; and
CC to isolate a full length gene from a suitable library e.g., a tumour cDNA
CC library using well known techniques
XX
SQ Sequence 144 BP; 61 A; 26 C; 46 G; 11 T; 0 U; 0 Other;

Query Match 2.0%; Score 22; DB 6; Length 144;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        862 AGCCTAAAAAAAGAAAAAAAAA 883
              |||||||||||||||||||||||
Db         99 AGCCTAAAAAAAGAAAAAAAAA 120
```

```
RESULT 4
ABL86321
ID   ABL86321 standard; cDNA; 154 BP.
XX
AC   ABL86321;
XX
DT   17-MAY-2002  (first entry)
XX
DE   Human ovarian cancer related cDNA clone SEQ ID NO:9299.
XX
KW   Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200192581-A2.
XX
PD   06-DEC-2001.
XX
PF   29-MAY-2001; 2001WO-US017756.
XX
PR   26-MAY-2000; 2000US-0207484P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Algate PA,  Harlocker SL,  Jones R;
XX
DR   WPI; 2002-122075/16.
XX
PT   Composition for therapy and diagnosis of ovarian cancer comprising
PT   polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT   polypeptide, antibody specific to polypeptide or T cell expressing
PT   polypeptide.
XX
PS   Claim 1; SEQ ID NO 9299; 489pp; English.
XX
CC   The present invention describes a composition (I) comprising: carriers
CC   and immunostimulants; and a polypeptide (II) of a ovarian tumour
CC   polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
CC   from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
CC   (III) encoding (II) having a sequence (S2), a T cell population of (II),
CC   or antigen presenting cells that express (II). (I) has cytostatic
CC   activity. An oligonucleotide (IV) that hybridises to (S1) can be used for
CC   detecting ovarian cancer in a patient's biological sample preferably
CC   serum or ovarian tissue. The method comprises contacting a biological
CC   sample from a patient with (IV), detecting the amount of polynucleotide
CC   hybridising to (IV) and comparing the amount to a predetermined cutoff
CC   value and thereby detecting ovarian cancer in the patient, where the
CC   amount of polynucleotide hybridising to (IV) is detected preferably by
CC   polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
CC   useful for stimulating and/or expanding T cells specific for an ovarian
CC   tumour protein comprising contacting T cells with (III) or (II). (III) is
CC   useful in design and preparation of ribozyme molecules for inhibiting
CC   expression of the tumour polypeptides and proteins in tumour cells; and
CC   to isolate a full length gene from a suitable library e.g., a tumour cDNA
CC   library using well known techniques
XX
SQ   Sequence 154 BP; 59 A; 28 C; 54 G; 13 T; 0 U; 0 Other;

  Query Match             2.0%;  Score 22;  DB 6;  Length 154;
  Best Local Similarity   100.0%;  Pred. No. 2;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        862 AGCCTAAAAAAAGAAAAAAAAA 883
              |||||||||||||||||||||
Db        112 AGCCTAAAAAAAGAAAAAAAAA 133


RESULT 5
ABL86948
ID   ABL86948 standard; cDNA; 213 BP.
XX
AC   ABL86948;
XX
DT   17-MAY-2002  (first entry)
XX
DE   Human ovarian cancer related cDNA clone SEQ ID NO:9926.
XX
KW   Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200192581-A2.
XX
PD   06-DEC-2001.
XX
PF   29-MAY-2001; 2001WO-US017756.
XX
PR   26-MAY-2000; 2000US-0207484P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Algate PA,  Harlocker SL,  Jones R;
XX
DR   WPI; 2002-122075/16.
XX
PT   Composition for therapy and diagnosis of ovarian cancer comprising
PT   polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT   polypeptide, antibody specific to polypeptide or T cell expressing
PT   polypeptide.
XX
PS   Claim 1; SEQ ID NO 9926; 489pp; English.
XX
CC   The present invention describes a composition (I) comprising: carriers
CC   and immunostimulants; and a polypeptide (II) of a ovarian tumour
CC   polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
CC   from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
CC   (III) encoding (II) having a sequence (S2), a T cell population of (II),
CC   or antigen presenting cells that express (II). (I) has cytostatic
CC   activity. An oligonucleotide (IV) that hybridises to (S1) can be used for
CC   detecting ovarian cancer in a patient's biological sample preferably
CC   serum or ovarian tissue. The method comprises contacting a biological
CC   sample from a patient with (IV), detecting the amount of polynucleotide
CC   hybridising to (IV) and comparing the amount to a predetermined cutoff
CC   value and thereby detecting ovarian cancer in the patient, where the
CC   amount of polynucleotide hybridising to (IV) is detected preferably by
CC   polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is
CC   useful for stimulating and/or expanding T cells specific for an ovarian
CC   tumour protein comprising contacting T cells with (III) or (II). (III) is
CC   useful in design and preparation of ribozyme molecules for inhibiting
CC   expression of the tumour polypeptides and proteins in tumour cells; and
CC   to isolate a full length gene from a suitable library e.g., a tumour cDNA
CC   library using well known techniques
XX
SQ   Sequence 213 BP; 79 A; 45 C; 70 G; 19 T; 0 U; 0 Other;

  Query Match             2.0%;  Score 22;  DB 6;  Length 213;
  Best Local Similarity   100.0%;  Pred. No. 2.1;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        862 AGCCTAAAAAAAGAAAAAAAAA 883
              |||||||||||||||||||||
Db        163 AGCCTAAAAAAAGAAAAAAAAA 184


RESULT 6
ADL62118/c
ID   ADL62118 standard; DNA; 1074 BP.
XX
AC   ADL62118;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Human ovarian cancer DNA marker #20330.
XX
```

KW Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.
XX
OS Homo sapiens.
XX
PN WO200170979-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US009126.
XX
PR 21-MAR-2000; 2000US-0191031P.
PR 25-MAY-2000; 2000US-0207124P.
PR 15-JUN-2000; 2000US-0211940P.
PR 07-JUL-2000; 2000US-0216820P.
PR 25-JUL-2000; 2000US-0220661P.
PR 21-DEC-2000; 2000US-0257672P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lee J, Lillie J;
XX
DR WPI; 2001-611502/70.
XX
PT Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
PT cancer cells as compared to their normal non-cancerous ovarian cells are
PT used to characterize stage, grade, histological type of ovarian cancer.
XX
PS Disclosure; SEQ ID NO 20330; 106pp; English.
XX
CC The invention relates to nucleic acid markers which are overexpressed in
CC ovarian cancer cells as compared to their expression in normal (i.e. non-
CC cancerous) ovarian cells. The invention also relates to polypeptides
CC encoded by the markers, antibodies that selectively bind to the
CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk
CC of developing ovarian cancer involving inhibiting expression of a gene
CC corresponding to a marker of the invention and a method of treating a
CC patient afflicted with ovarian cancer comprising providing to cells of
CC the patient an antisense oligonucleotide complementary to a marker of the
CC invention. The markers are useful for assessing if a patient is afflicted
CC with ovarian cancer, which involves comparing the level of expression of
CC a marker in a patient sample and a normal level of expression of the
CC marker in a control non-ovarian cancer sample. A difference between the
CC expression levels indicates ovarian cancer. The level of expression of a
CC marker corresponds to a secreted protein or to a transcribed
CC polynucleotide or its portion. The level of expression of the marker is
CC assessed by detecting the presence in the sample, a protein or protein
CC fragment corresponding to the marker. The presence of protein or protein
CC fragment is detected using an antibody that specifically binds with the
CC protein or protein fragment. Alternatively, the level of expression of
CC the marker is assessed by detecting the presence of a transcribed
CC polynucleotide which anneals with the marker or anneals with a portion of
CC the polynucleotide comprising the marker, under stringent conditions. The
CC marker is also used for monitoring the progression of ovarian cancer in a
CC patient which involves detecting expression of the marker in a patient
CC sample at a first point in time, repeating the method at a subsequent
CC time and comparing the level of expression. The method is carried out
CC using an ovarian tissue sample. A composition comprising a marker,
CC polypeptide or antibody of the invention is used to treat ovarian cancer.
CC This sequence represents a human ovarian cancer DNA marker of the
CC invention.
XX
SQ Sequence 1074 BP; 194 A; 278 C; 290 G; 300 T; 0 U; 12 Other;

Query Match 2.0%; Score 22; DB 5; Length 1074;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 867 AAAAAAAGAAAAAAAAACTGTT 888
||||||||||||||||||||||
Db 427 AAAAAAAGAAAAAAAAACTGTT 406

RESULT 7
ABL10390/c
ID ABL10390 standard; cDNA; 24935 BP.
XX
AC ABL10390;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 25652.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR P-PSDB; ABB66287.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 25652; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 24935 BP; 7255 A; 5321 C; 5280 G; 7079 T; 0 U; 0 Other;

Query Match 2.0%; Score 22; DB 4; Length 24935;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 872 AAGAAAAAAAAACTGTTCAATC 893
||||||||||||||||||||||
Db 12697 AAGAAAAAAAAACTGTTCAATC 12676

RESULT 8
AAV86132/c
ID AAV86132 standard; cDNA; 377 BP.
XX
AC AAV86132;
XX
DT 27-APR-1999 (first entry)
XX
DE EST clone H165.
XX
KW Expressed sequence tag; secreted protein; haematopoiesis regulator;
KW tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
XX

OS Homo sapiens.
XX
PN WO9845435-A2.
XX
PD 15-OCT-1998.
XX
PF 10-APR-1998; 98WO-US006954.
XX
PR 10-APR-1997; 97US-00835913.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Jacobs K, Mccoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;
PI Spaulding V, Agostino MJ;
XX
DR WPI; 1999-070076/06.
XX
PT New polynucleotides encoding human secreted proteins - derived from e.g.
PT human blood, kidney, foetal lung, placenta, testes, brain, ovary,
PT pituitary, retina and colon cDNA libraries.
XX
PS Claim 1; Page 132; 633pp; English.
XX
CC This sequence represents an expressed sequence tag (EST), and is a
CC polynucleotide of the invention. The polynucleotides of the invention are
CC all secreted EST sequences isolated from a variety of human tissue
CC sources. The EST sequences and proteins encoded by them are predicted to
CC have useful biological activities which would make them suitable for
CC treating, preventing or ameliorating medical conditions in humans and
CC animals, although no supporting data is given. Suggested activities
CC include nutritional activity, immune stimulating or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
CC activity. The EST sequences are also stated to be useful for gene therapy
XX
SQ Sequence 377 BP; 89 A; 88 C; 76 G; 124 T; 0 U; 0 Other;

Query Match 1.9%; Score 21; DB 2; Length 377;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 863 GCCTAAAAAAAGAAAAAAAAA 883
||||||||||||||||||||
Db 43 GCCTAAAAAAAGAAAAAAAAA 23

RESULT 9
AAC24900
ID AAC24900 standard; cDNA; 403 BP.
XX
AC AAC24900;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 28975.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-00200610.
XX
PR 26-FEB-1999; 99US-0122487P.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS Claim 1; SEQ ID NO 28975; 71pp + Sequence Listing; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors
XX
SQ Sequence 403 BP; 141 A; 68 C; 60 G; 128 T; 0 U; 6 Other;

Query Match 1.9%; Score 21; DB 3; Length 403;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 863 GCCTAAAAAAAGAAAAAAAAA 883
||||||||||||||||||||
Db 163 GCCTAAAAAAAGAAAAAAAAA 183

RESULT 10
AAH71551
ID AAH71551 standard; cDNA; 621 BP.
XX
AC AAH71551;
XX
DT 19-SEP-2001 (first entry)
XX
DE Human cervical cancer marker nucleic acid 2825.
XX
KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200142467-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-US033312.
XX
PR 08-DEC-1999; 99US-0169681P.
PR 21-DEC-1999; 99US-0171350P.
PR 14-MAR-2000; 2000US-0189315P.
PR 12-MAY-2000; 2000US-0203791P.
PR 09-JUN-2000; 2000US-0210600P.
PR 21-JUL-2000; 2000US-0220114P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Deeds J, Berger A, Zhao X;
XX
DR WPI; 2001-375006/39.
XX
PT New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.
XX
PS Claim 1; Page 579; 1051pp; English.

```
XX
CC   The invention relates to novel genes (AAH68727-AAH73383) associated with
CC   cervical cancer with cytostatic activity. The nucleic acids and encoded
CC   polypeptides are useful: to assess if a patient is afflicted with
CC   cervical cancer or has a pre-malignant condition; to monitor the
CC   progression of cervical cancer or a premalignant condition in a patient;
CC   and to select and/or assess the efficacy of a compound or therapy for
CC   inhibiting cervical cancer in a patient. The nucleic acids may also be
CC   useful for gene therapy
XX
SQ   Sequence 621 BP; 273 A; 63 C; 47 G; 238 T; 0 U; 0 Other;

  Query Match             1.9%;  Score 21;  DB 4;  Length 621;
  Best Local Similarity   100.0%;  Pred. No. 6.9;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        864 CCTAAAAAAAGAAAAAAAAAC 884
              |||||||||||||||||||||
Db        601 CCTAAAAAAAGAAAAAAAAAC 621


RESULT 11
AAF16195/c
ID   AAF16195 standard; cDNA; 1097 BP.
XX
AC   AAF16195;
XX
DT   13-MAR-2001  (first entry)
XX
DE   Human prostate cancer antigen nucleotide sequence SEQ ID NO:630.
XX
KW   Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW   neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW   vulnerary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW   antibacterial; gene therapy; neural; immune; reproductive; renal;
KW   gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW   wound; infectious disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055174-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US005988.
XX
PR   12-MAR-1999;   99US-0124270P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (ROSE/) ROSEN C A.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587513/55.
DR   P-PSDB; AAB56992.
XX
PT   Prostate cancer associated gene sequences, referred to as prostate cancer
PT   antigens, useful for treatment, prevention, and diagnosis of disorders
PT   such as prostate cancer.
XX
PS   Claim 1; Page 1072-1073; 2338pp; English.
XX
CC   AAF15566 to AAF16505 encode the human prostate cancer associated
CC   proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC   The prostate cancer antigens can have neuroprotective, cytostatic,
CC   cardioactive, immunomodulatory, muscular, vulnerary, gastrointestinal,
CC   nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC   and can be used in gene therapy. The prostate cancer antigen
CC   polynucleotides may be used for detection of prostate cancer, chromosome
CC   identification, as chromosome markers, and for numerous other diagnostic
CC   or research purposes. The prostate cancer antigens may be used to treat
CC   disorders such as neural, immune, muscular, reproductive,
CC   gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC   disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC   AAB57303 represent sequences used in the exemplification of the present
CC   invention
XX
SQ   Sequence 1097 BP; 301 A; 204 C; 201 G; 389 T; 0 U; 2 Other;

  Query Match             1.9%;  Score 21;  DB 3;  Length 1097;
  Best Local Similarity   100.0%;  Pred. No. 7;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        863 GCCTAAAAAAAGAAAAAAAAA 883
              |||||||||||||||||||||
Db        162 GCCTAAAAAAAGAAAAAAAAA 142


RESULT 12
ADA72818/c
ID   ADA72818 standard; DNA; 2000 BP.
XX
AC   ADA72818;
XX
DT   20-NOV-2003  (first entry)
XX
DE   Rice gene, SEQ ID 6143.
XX
KW   Plant; bacterial infection; fungal infection; viral infection; rice;
KW   gene; ds.
XX
OS   Oryza sativa.
XX
PN   WO2003000898-A1.
XX
PD   03-JAN-2003.
XX
PF   22-JUN-2001; 2001WO-IB001105.
XX
PR   22-JUN-2001; 2001WO-IB001105.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Chang H,  Chen W,  Cooper B,  Glazebrook J,  Goff SA,  Hou Y;
PI   Katagiri F,  Quan S,  Tao Y,  Whitham S,  Xie Z,  Zhu T,  Zou G;
XX
DR   WPI; 2003-175290/17.
XX
PT   Identifying at least one gene involved in plant resistance or response to
PT   pathogenic infection for conferring resistance or tolerance to a plant to
PT   bacterial, fungal or viral infection by determining or detecting plant
PT   gene expression.
XX
PS   Claim 27; SEQ ID NO 6143; 899pp; English.
XX
CC   The present invention relates to a method (M1) for identifying genes
CC   involved in plant resistance or response to pathogenic infection. M1
CC   comprises identifying a gene whose expression is significantly altered in
CC   the incompatible interaction of plant gene expression relative to
CC   expression of the gene in an uninfected plant, in a mutant plant that
CC   does not express a gene associated with response to pathogenic infection,
CC   or in a corresponding incompatible or compatible interaction. (M1) is
CC   useful for conferring resistance to resistance or tolerance to a plant to
CC   bacterial, fungal or viral infection. The present sequence was used to
CC   illustrate the invention.
XX
SQ   Sequence 2000 BP; 632 A; 370 C; 326 G; 672 T; 0 U; 0 Other;

  Query Match             1.9%;  Score 21;  DB 8;  Length 2000;
  Best Local Similarity   100.0%;  Pred. No. 7.2;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        867 AAAAAAAGAAAAAAAAACTGT 887
              |||||||||||||||||||||
Db       1385 AAAAAAAGAAAAAAAAACTGT 1365
```

RESULT 13
AAH72880/c
ID AAH72880 standard; cDNA; 2170 BP.
XX
AC AAH72880;
XX
DT 19-SEP-2001 (first entry)
XX
DE Human cervical cancer marker nucleic acid 4154.
XX
KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200142467-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-US033312.
XX
PR 08-DEC-1999; 99US-0169681P.
PR 21-DEC-1999; 99US-0171350P.
PR 14-MAR-2000; 2000US-0189315P.
PR 12-MAY-2000; 2000US-0203791P.
PR 09-JUN-2000; 2000US-0210600P.
PR 21-JUL-2000; 2000US-0220114P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Deeds J, Berger A, Zhao X;
XX
DR WPI; 2001-375006/39.
XX
PT New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.
XX
PS Claim 1; Page 903; 1051pp; English.
XX
CC The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy
XX
SQ Sequence 2170 BP; 358 A; 689 C; 427 G; 688 T; 0 U; 8 Other;

Query Match 1.9%; Score 21; DB 4; Length 2170;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 866 TAAAAAAAGAAAAAAAAACTG 886
||||||||||||||||||||
Db 96 TAAAAAAAGAAAAAAAAACTG 76

RESULT 14
ABV22589/c
ID ABV22589 standard; cDNA; 2170 BP.
XX
AC ABV22589;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 22580.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 3950; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodyanamic or pharmacogenomic marker
XX
SQ Sequence 2170 BP; 358 A; 688 C; 428 G; 688 T; 0 U; 8 Other;

Query Match 1.9%; Score 21; DB 5; Length 2170;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 866 TAAAAAAAGAAAAAAAAACTG 886
||||||||||||||||||||
Db 96 TAAAAAAAGAAAAAAAAACTG 76

RESULT 15
ABV25695/c
ID ABV25695 standard; cDNA; 2170 BP.
XX
AC ABV25695;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 25686.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.

PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, · Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 5153; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodyanamic or pharmacogenomic marker
XX
SQ Sequence 2170 BP; 358 A; 688 C; 428 G; 688 T; 0 U; 8 Other;

Query Match 1.9%; Score 21; DB 5; Length 2170;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 866 TAAAAAAAGAAAAAAAAACTG 886
||||||||||||||||||||
Db 96 TAAAAAAAGAAAAAAAAACTG 76

Search completed: August 29, 2005, 03:26:35
Job time : 702 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 29, 2005, 03:00:28 ; Search time 233 Seconds
(without alignments)
7802.161 Million cell updates/sec

Title: US-09-979-549-2
Perfect score: 1111
Sequence: 1 gaagtcagaaggccgttcag.....ccgagaagagagagagagag 1111

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	22	2.0	249	4	US-09-248-796A-9416	Sequence 9416, Ap
c	2	22	2.0	601	4	US-09-949-016-119414	Sequence 119414,
c	3	22	2.0	601	4	US-09-949-016-119415	Sequence 119415,
c	4	22	2.0	51403	4	US-09-949-016-15057	Sequence 15057, A
c	5	22	2.0	250352	4	US-09-949-016-14724	Sequence 14724, A
	6	21	1.9	403	4	US-09-513-999C-28975	Sequence 28975, A
	7	21	1.9	601	4	US-09-949-016-120940	Sequence 120940,
c	8	21	1.9	601	4	US-09-949-016-169986	Sequence 169986,
c	9	21	1.9	33908	4	US-09-949-016-15104	Sequence 15104, A
	10	21	1.9	114139	4	US-09-949-016-16536	Sequence 16536, A
c	11	20	1.8	601	4	US-09-949-016-80010	Sequence 80010, A
c	12	20	1.8	601	4	US-09-949-016-80011	Sequence 80011, A
c	13	20	1.8	601	4	US-09-949-016-80012	Sequence 80012, A
c	14	20	1.8	601	4	US-09-949-016-136288	Sequence 136288,
c	15	20	1.8	601	4	US-09-949-016-148215	Sequence 148215,
c	16	20	1.8	601	4	US-09-949-016-148216	Sequence 148216,
c	17	20	1.8	667	4	US-09-771-035A-18	Sequence 18, Appl
	18	20	1.8	1725	4	US-09-949-016-3377	Sequence 3377, Ap
	19	20	1.8	1890	4	US-09-949-016-4824	Sequence 4824, Ap
	20	20	1.8	8625	4	US-09-949-016-16566	Sequence 16566, A
	21	20	1.8	13985	4	US-09-949-016-15640	Sequence 15640, A
	22	20	1.8	15632	4	US-09-949-016-15119	Sequence 15119, A
c	23	20	1.8	22471	4	US-09-949-016-15902	Sequence 15902, A
	24	20	1.8	91772	4	US-09-949-016-15568	Sequence 15568, A
	25	20	1.8	154023	4	US-09-949-016-17057	Sequence 17057, A
c	26	20	1.8	171700	4	US-09-949-016-12276	Sequence 12276, A
c	27	20	1.8	171701	4	US-09-949-016-15835	Sequence 15835, A
c	28	20	1.8	312470	4	US-09-949-016-14043	Sequence 14043, A
c	29	20	1.8	336024	4	US-09-949-016-12373	Sequence 12373, A
c	30	20	1.8	636591	4	US-09-949-016-11808	Sequence 11808, A
c	31	20	1.8	636591	4	US-09-949-016-13388	Sequence 13388, A
c	32	19	1.7	375	4	US-09-621-976-13110	Sequence 13110, A
	33	19	1.7	489	4	US-09-252-991A-5369	Sequence 5369, Ap
c	34	19	1.7	555	3	US-08-479-089A-4	Sequence 4, Appli
c	35	19	1.7	555	4	US-07-669-545B-4	Sequence 4, Appli
	36	19	1.7	601	4	US-09-949-016-60407	Sequence 60407, A
	37	19	1.7	601	4	US-09-949-016-68221	Sequence 68221, A
	38	19	1.7	601	4	US-09-949-016-68222	Sequence 68222, A
	39	19	1.7	601	4	US-09-949-016-74691	Sequence 74691, A
	40	19	1.7	601	4	US-09-949-016-74692	Sequence 74692, A
	41	19	1.7	601	4	US-09-949-016-74693	Sequence 74693, A
	42	19	1.7	601	4	US-09-949-016-74694	Sequence 74694, A
	43	19	1.7	601	4	US-09-949-016-74890	Sequence 74890, A
	44	19	1.7	601	4	US-09-949-016-74891	Sequence 74891, A
	45	19	1.7	601	4	US-09-949-016-74892	Sequence 74892, A

ALIGNMENTS

RESULT 1
US-09-248-796A-9416/c
; Sequence 9416, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 9416
; LENGTH: 249
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-9416

Query Match 2.0%; Score 22; DB 4; Length 249;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        864 CCTAAAAAAAGAAAAAAAAACT 885
              ||||||||||||||||||||||
Db        108 CCTAAAAAAAGAAAAAAAAACT 87
```

RESULT 2
US-09-949-016-119414/c
; Sequence 119414, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0

```
; SEQ ID NO 119414
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-119414

  Query Match             2.0%;  Score 22;  DB 4;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 0.77;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        868 AAAAAAGAAAAAAAAACTGTTC 889
              |||||||||||||||||||||
Db        378 AAAAAAGAAAAAAAAACTGTTC 357


RESULT 3
US-09-949-016-119415/c
; Sequence 119415, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 119415
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-119415

  Query Match             2.0%;  Score 22;  DB 4;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 0.77;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        868 AAAAAAGAAAAAAAAACTGTTC 889
              |||||||||||||||||||||
Db         60 AAAAAAGAAAAAAAAACTGTTC 39


RESULT 4
US-09-949-016-15057/c
; Sequence 15057, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15057
;   LENGTH: 51403
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(51403)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15057

  Query Match             2.0%;  Score 22;  DB 4;  Length 51403;
  Best Local Similarity   100.0%;  Pred. No. 0.82;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        868 AAAAAAGAAAAAAAAACTGTTC 889
              |||||||||||||||||||||
Db      47891 AAAAAAGAAAAAAAAACTGTTC 47870


RESULT 5
US-09-949-016-14724/c
; Sequence 14724, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14724
;   LENGTH: 250352
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(250352)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14724

  Query Match             2.0%;  Score 22;  DB 4;  Length 250352;
  Best Local Similarity   100.0%;  Pred. No. 0.83;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        865 CTAAAAAAAGAAAAAAAAACTG 886
              |||||||||||||||||||||
Db     228042 CTAAAAAAAGAAAAAAAAACTG 228021


RESULT 6
US-09-513-999C-28975
; Sequence 28975, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 28975
;   LENGTH: 403
;   TYPE: DNA
;   ORGANISM: Homo sapiens
```

```
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 56
;   OTHER INFORMATION: r=a or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 254
;   OTHER INFORMATION: w=a or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 255
;   OTHER INFORMATION: n=a, g, c or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 271
;   OTHER INFORMATION: n=a, g, c or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 332
;   OTHER INFORMATION: h=a or c or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 358
;   OTHER INFORMATION: v=a or c or g
US-09-513-999C-28975

  Query Match             1.9%;  Score 21;  DB 4;  Length 403;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        863 GCCTAAAAAAAGAAAAAAAAA 883
              |||||||||||||||||||||
Db        163 GCCTAAAAAAAGAAAAAAAAA 183


RESULT 7
US-09-949-016-120940
; Sequence 120940, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 120940
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-120940

  Query Match             1.9%;  Score 21;  DB 4;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        867 AAAAAAAGAAAAAAAAACTGT 887
              |||||||||||||||||||||
Db        401 AAAAAAAGAAAAAAAAACTGT 421


RESULT 8
US-09-949-016-169986/c
; Sequence 169986, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 169986
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-169986

  Query Match             1.9%;  Score 21;  DB 4;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        868 AAAAAAGAAAAAAAAACTGTT 888
              |||||||||||||||||||||
Db        426 AAAAAAGAAAAAAAAACTGTT 406


RESULT 9
US-09-949-016-15104/c
; Sequence 15104, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15104
;   LENGTH: 33908
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(33908)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15104

  Query Match             1.9%;  Score 21;  DB 4;  Length 33908;
  Best Local Similarity   100.0%;  Pred. No. 2.6;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        867 AAAAAAAGAAAAAAAAACTGT 887
              |||||||||||||||||||||
Db      29559 AAAAAAAGAAAAAAAAACTGT 29539


RESULT 10
US-09-949-016-16536
; Sequence 16536, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16536
;   LENGTH: 114139
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(114139)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16536

  Query Match             1.9%;  Score 21;  DB 4;  Length 114139;
  Best Local Similarity   100.0%;  Pred. No. 2.6;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        868 AAAAAAGAAAAAAAAACTGTT 888
              |||||||||||||||||||||
Db      57687 AAAAAAGAAAAAAAAACTGTT 57707


RESULT 11
US-09-949-016-80010/c
; Sequence 80010, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80010
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-80010

  Query Match             1.8%;  Score 20;  DB 4;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 7.9;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        864 CCTAAAAAAAGAAAAAAAAA 883
              ||||||||||||||||||||
Db        368 CCTAAAAAAAGAAAAAAAAA 349


RESULT 12
US-09-949-016-80011/c
; Sequence 80011, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80011
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-80011

  Query Match             1.8%;  Score 20;  DB 4;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 7.9;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        864 CCTAAAAAAAGAAAAAAAAA 883
              ||||||||||||||||||||
Db        321 CCTAAAAAAAGAAAAAAAAA 302


RESULT 13
US-09-949-016-80012/c
; Sequence 80012, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 80012
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-80012

  Query Match             1.8%;  Score 20;  DB 4;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 7.9;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        864 CCTAAAAAAAGAAAAAAAAA 883
              ||||||||||||||||||||
Db        220 CCTAAAAAAAGAAAAAAAAA 201


RESULT 14
US-09-949-016-136288/c
; Sequence 136288, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
```

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 136288
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-136288

Query Match 1.8%; Score 20; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 875 AAAAAAAAACTGTTCAATCA 894
||||||||||||||||||||
Db 309 AAAAAAAAACTGTTCAATCA 290

RESULT 15
US-09-949-016-148215/c
; Sequence 148215, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 148215
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-148215

Query Match 1.8%; Score 20; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 416 GGAGGAGAAAGAAAGTAGGA 435
||||||||||||||||||||
Db 335 GGAGGAGAAAGAAAGTAGGA 316

Search completed: August 29, 2005, 06:07:46
Job time : 236 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 29, 2005, 03:14:53 ; Search time 2629 Seconds
(without alignments)
2765.071 Million cell updates/sec

Title: US-09-979-549-2
Perfect score: 1111
Sequence: 1 gaagtcagaaggccgttcag.....ccgagaagagagagagagag 1111

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 7331713 seqs, 3271544945 residues

Word size : 0

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	42	3.8	629	20	US-10-425-115-15660	Sequence 15660, A
2	22	2.0	141	9	US-09-867-701-9430	Sequence 9430, Ap
3	22	2.0	144	9	US-09-867-701-9165	Sequence 9165, Ap
4	22	2.0	154	9	US-09-867-701-9299	Sequence 9299, Ap
5	22	2.0	213	9	US-09-867-701-9926	Sequence 9926, Ap
c 6	22	2.0	1074	10	US-09-814-353-20330	Sequence 20330, A
c 7	22	2.0	394468	21	US-10-741-600-17952	Sequence 17952, A
8	21	1.9	261	17	US-10-242-535A-46223	Sequence 46223, A
9	21	1.9	261	18	US-10-085-783A-46223	Sequence 46223, A
c 10	21	1.9	542	20	US-10-425-115-101740	Sequence 101740,
c 11	21	1.9	614	19	US-10-437-963-3058	Sequence 3058, Ap
12	21	1.9	828	13	US-10-027-632-169396	Sequence 169396,
13	21	1.9	828	17	US-10-027-632-169396	Sequence 169396,
c 14	21	1.9	1097	9	US-09-925-300-630	Sequence 630, App
15	21	1.9	1420	19	US-10-437-963-85377	Sequence 85377, A
c 16	21	1.9	2170	20	US-10-357-930-22582	Sequence 22582, A
c 17	21	1.9	2170	20	US-10-357-930-24738	Sequence 24738, A
c 18	21	1.9	2170	20	US-10-357-930-25684	Sequence 25684, A
c 19	21	1.9	2170	20	US-10-357-930-28422	Sequence 28422, A
c 20	21	1.9	3191	10	US-09-374-046A-59	Sequence 59, Appl
c 21	21	1.9	3191	18	US-10-616-263-59	Sequence 59, Appl
c 22	21	1.9	3618	18	US-10-380-374-21	Sequence 21, Appl
c 23	21	1.9	4740	20	US-10-357-930-21715	Sequence 21715, A
c 24	21	1.9	4740	20	US-10-357-930-27560	Sequence 27560, A
c 25	21	1.9	10351	9	US-09-874-470-5	Sequence 5, Appli
c 26	21	1.9	80815	19	US-10-322-281-486	Sequence 486, App
27	20	1.8	162	17	US-10-131-827-8661	Sequence 8661, Ap
28	20	1.8	369	13	US-10-027-632-68718	Sequence 68718, A
29	20	1.8	369	17	US-10-027-632-68718	Sequence 68718, A
c 30	20	1.8	371	10	US-09-960-706-561	Sequence 561, App
c 31	20	1.8	371	10	US-09-873-319-346	Sequence 346, App
32	20	1.8	375	20	US-10-357-930-48707	Sequence 48707, A
33	20	1.8	422	11	US-09-969-034-2636	Sequence 2636, Ap
34	20	1.8	462	10	US-09-918-995-9917	Sequence 9917, Ap
35	20	1.8	555	9	US-09-764-869-2246	Sequence 2246, Ap
36	20	1.8	555	14	US-10-091-504-2246	Sequence 2246, Ap
37	20	1.8	555	17	US-10-227-577-2246	Sequence 2246, Ap
38	20	1.8	571	13	US-10-027-632-47881	Sequence 47881, A
39	20	1.8	571	13	US-10-027-632-295267	Sequence 295267,
40	20	1.8	571	17	US-10-027-632-47881	Sequence 47881, A
41	20	1.8	571	17	US-10-027-632-295267	Sequence 295267,
c 42	20	1.8	630	13	US-10-027-632-280598	Sequence 280598,
c 43	20	1.8	630	17	US-10-027-632-280598	Sequence 280598,
c 44	20	1.8	667	9	US-09-771-035A-18	Sequence 18, Appl
c 45	20	1.8	667	18	US-10-634-221-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-10-425-115-15660
; Sequence 15660, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 15660
; LENGTH: 629
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_114281C.1
US-10-425-115-15660

Query Match 3.8%; Score 42; DB 20; Length 629;
Best Local Similarity 100.0%; Pred. No. 5.3e-12;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1033 GAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAA 1074
||||||||||||||||||||||||||||||||||||||||||
Db 26 GAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAA 67

RESULT 2
US-09-867-701-9430
; Sequence 9430, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9430
; LENGTH: 141
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-9430

Query Match 2.0%; Score 22; DB 9; Length 141;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 862 AGCCTAAAAAAAGAAAAAAAAA 883
||||||||||||||||||||||
Db 98 AGCCTAAAAAAAGAAAAAAAAA 119

RESULT 3
US-09-867-701-9165
; Sequence 9165, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9165
; LENGTH: 144
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-9165

Query Match 2.0%; Score 22; DB 9; Length 144;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 862 AGCCTAAAAAAAGAAAAAAAAA 883
||||||||||||||||||||||
Db 99 AGCCTAAAAAAAGAAAAAAAAA 120

RESULT 4
US-09-867-701-9299
; Sequence 9299, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9299
; LENGTH: 154
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-9299

Query Match 2.0%; Score 22; DB 9; Length 154;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 862 AGCCTAAAAAAAGAAAAAAAAA 883
||||||||||||||||||||||
Db 112 AGCCTAAAAAAAGAAAAAAAAA 133

RESULT 5
US-09-867-701-9926
; Sequence 9926, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9926
; LENGTH: 213
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-9926

Query Match 2.0%; Score 22; DB 9; Length 213;
Best Local Similarity 100.0%; Pred. No. 0.63;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 862 AGCCTAAAAAAAGAAAAAAAAA 883
||||||||||||||||||||||
Db 163 AGCCTAAAAAAAGAAAAAAAAA 184

RESULT 6
US-09-814-353-20330/c
; Sequence 20330, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25

```
;  PRIOR APPLICATION NUMBER: US 60/257,672
;  PRIOR FILING DATE: 2000-12-21
;  NUMBER OF SEQ ID NOS: 22037
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20330
;   LENGTH: 1074
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 968, 1074
;   OTHER INFORMATION: n = A,T,C or G
US-09-814-353-20330

  Query Match             2.0%;  Score 22;  DB 10;  Length 1074;
  Best Local Similarity   100.0%;  Pred. No. 0.69;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        867 AAAAAAAGAAAAAAAAACTGTT 888
              ||||||||||||||||||||||
Db        427 AAAAAAAGAAAAAAAAACTGTT 406


RESULT 7
US-10-741-600-17952/c
; Sequence 17952, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
;  APPLICANT: CARGILL, Michele et al.
;  TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
;  TITLE OF INVENTION:  MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001499
;  CURRENT APPLICATION NUMBER: US/10/741,600
;  CURRENT FILING DATE:  2003-12-22
;  NUMBER OF SEQ ID NOS: 73997
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17952
;   LENGTH: 394468
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(394468)
;   OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-600-17952

  Query Match             2.0%;  Score 22;  DB 21;  Length 394468;
  Best Local Similarity   100.0%;  Pred. No. 0.92;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        865 CTAAAAAAAGAAAAAAAAACTG 886
              ||||||||||||||||||||||
Db      53760 CTAAAAAAAGAAAAAAAAACTG 53739


RESULT 8
US-10-242-535A-46223
; Sequence 46223, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 46223
;   LENGTH: 261
;   TYPE: DNA
;   ORGANISM: Human
US-10-242-535A-46223

  Query Match             1.9%;  Score 21;  DB 17;  Length 261;
  Best Local Similarity   100.0%;  Pred. No. 2.3;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        863 GCCTAAAAAAAGAAAAAAAAA 883
              |||||||||||||||||||||
Db        222 GCCTAAAAAAAGAAAAAAAAA 242


RESULT 9
US-10-085-783A-46223
; Sequence 46223, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2002
;  CURRENT APPLICATION NUMBER: US/10/085,783A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 46223
;   LENGTH: 261
;   TYPE: DNA
;   ORGANISM: Human
US-10-085-783A-46223

  Query Match             1.9%;  Score 21;  DB 18;  Length 261;
  Best Local Similarity   100.0%;  Pred. No. 2.3;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        863 GCCTAAAAAAAGAAAAAAAAA 883
              |||||||||||||||||||||
Db        222 GCCTAAAAAAAGAAAAAAAAA 242


RESULT 10
US-10-425-115-101740/c
; Sequence 101740, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 101740
;   LENGTH: 542
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_24297C.1
```

US-10-425-115-101740

Query Match 1.9%; Score 21; DB 20; Length 542;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 863 GCCTAAAAAAAGAAAAAAAAA 883
||||||||||||||||||||
Db 50 GCCTAAAAAAAGAAAAAAAAA 30

RESULT 11
US-10-437-963-3058/c
; Sequence 3058, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 3058
; LENGTH: 614
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_102776C.1
US-10-437-963-3058

Query Match 1.9%; Score 21; DB 19; Length 614;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1091 GCCGAGAAGAGAGAGAGAGAG 1111
||||||||||||||||||||
Db 108 GCCGAGAAGAGAGAGAGAGAG 88

RESULT 12
US-10-027-632-169396
; Sequence 169396, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 169396
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-169396

Query Match 1.9%; Score 21; DB 13; Length 828;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 867 AAAAAAAGAAAAAAAAACTGT 887
||||||||||||||||||||
Db 118 AAAAAAAGAAAAAAAAACTGT 138

RESULT 13
US-10-027-632-169396
; Sequence 169396, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 169396
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-169396

Query Match 1.9%; Score 21; DB 17; Length 828;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 867 AAAAAAAGAAAAAAAAACTGT 887
||||||||||||||||||||
Db 118 AAAAAAAGAAAAAAAAACTGT 138

RESULT 14
US-09-925-300-630/c
; Sequence 630, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 630
; LENGTH: 1097
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (13)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-630

Query Match 1.9%; Score 21; DB 9; Length 1097;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 863 GCCTAAAAAAAGAAAAAAAAA 883
||||||||||||||||||||
Db 162 GCCTAAAAAAAGAAAAAAAAA 142

RESULT 15
US-10-437-963-85377
; Sequence 85377, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 85377
; LENGTH: 1420
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1420)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_84521C.1
US-10-437-963-85377

Query Match 1.9%; Score 21; DB 19; Length 1420;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 870 AAAAGAAAAAAAAACTGTTCA 890
||||||||||||||||||||
Db 1363 AAAAGAAAAAAAAACTGTTCA 1383

Search completed: August 29, 2005, 06:51:41
Job time : 2631 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 29, 2005, 02:56:28 ; Search time 4261 Seconds
(without alignments)
9924.760 Million cell updates/sec

Title: US-09-979-549-2
Perfect score: 1111
Sequence: 1 gaagtcagaaggccgttcag.....ccgagaagagagagagagag 1111

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		99	8.9	509	6	CD867573	CD867573 AZO2.106K
2		99	8.9	570	6	CD867174	CD867174 AZO2.105J
3		92	8.3	594	6	CA593541	CA593541 wpa1c.pk0
4		92	8.3	631	6	CD938039	CD938039 OV.108O07
5		90	8.1	624	6	CD865240	CD865240 AZO2.073I
6		88	7.9	430	6	CA712930	CA712930 wdk3c.pk0
7		66	5.9	643	6	CD884411	CD884411 F1.116I20
8		64	5.8	642	7	CF132917	CF132917 WHE4351_G
9		64	5.8	712	6	CD894217	CD894217 G118.125L
10		63	5.7	556	6	CA701748	CA701748 wkm2c.pk0
11		63	5.7	637	6	CD904051	CD904051 G356.112E
12		62	5.6	422	2	BE419251	BE419251 WWR021.H3
13		62	5.6	578	4	BJ257084	BJ257084 BJ257084
c 14		62	5.6	579	4	BJ221033	BJ221033 BJ221033
c 15		60	5.4	597	4	BJ253061	BJ253061 BJ253061
c 16		58	5.2	533	4	BJ239800	BJ239800 BJ239800
17		58	5.2	585	4	BJ233935	BJ233935 BJ233935
18		56	5.0	543	6	CA497278	CA497278 WHE3225_C
c 19		56	5.0	1061	7	CK213031	CK213031 FGAS02493
c 20		55	5.0	593	4	BJ296612	BJ296612 BJ296612
21		55	5.0	605	4	BJ290275	BJ290275 BJ290275
c 22		54	4.9	582	4	BJ262654	BJ262654 BJ262654
c 23		54	4.9	875	7	CK159678	CK159678 FGAS04115
c 24		53	4.8	839	7	CK158306	CK158306 FGAS03954
c 25		53	4.8	889	7	CK158740	CK158740 FGAS04003
26		52	4.7	418	6	CA702286	CA702286 wdk1c.pk0
27		47	4.2	557	4	BJ223826	BJ223826 BJ223826
28		47	4.2	628	5	BQ805515	BQ805515 WHE3567_H
29		44	4.0	353	2	BE406240	BE406240 WHE0409_D
c 30		43	3.9	563	4	BJ273554	BJ273554 BJ273554
31		43	3.9	565	5	BU101004	BU101004 WHE3360_F
32		43	3.9	567	6	CA717478	CA717478 wdk4c.pk0
33		43	3.9	614	6	CA597581	CA597581 wpa1c.pk0
c 34		43	3.9	630	4	BJ320546	BJ320546 BJ320546
35		43	3.9	630	6	CD930486	CD930486 GR45.111H
36		43	3.9	644	6	CD874126	CD874126 AZO3.101H
c 37		43	3.9	1141	7	CK211149	CK211149 FGAS02298
38		41	3.7	459	1	AL826362	AL826362 AL826362
39		40	3.6	156	2	BE419352	BE419352 WWS01.A5R
40		40	3.6	608	4	BJ223851	BJ223851 BJ223851
41		40	3.6	633	4	BJ258073	BJ258073 BJ258073
42		34	3.1	577	4	BJ214045	BJ214045 BJ214045
c 43		34	3.1	887	7	CK158744	CK158744 FGAS04003
44		32	2.9	137	4	BJ223280	BJ223280 BJ223280
45		32	2.9	160	4	BJ290875	BJ290875 BJ290875

ALIGNMENTS

RESULT 1
CD867573
LOCUS CD867573 509 bp mRNA linear EST 11-JUL-2003
DEFINITION AZO2.106K05F001108 AZO2 Triticum aestivum cDNA clone AZO2106K05, mRNA sequence.
ACCESSION CD867573
VERSION CD867573.1 GI:32551389
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 509)
AUTHORS Genoplante.
TITLE Genoplante, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french plant genomics programme 'Genoplante' (http://www.genoplante.com and http://genoplante-info.infobiogen.fr).
FEATURES Location/Qualifiers
source 1. .509
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="AZO2106K05"
/tissue_type="root"
/clone_lib="AZO2"
ORIGIN

Query Match 8.9%; Score 99; DB 6; Length 509;
Best Local Similarity 100.0%; Pred. No. 1.4e-41;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1013 GGGGAGGCGGGCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGA 1072
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGGAGGCGGGCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGA 60

Qy       1073 AATCAACAAGCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 1111
             |||||||||||||||||||||||||||||||||||||||||
Db         61 AATCAACAAGCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 99
```

```
RESULT 2
CD867174
LOCUS       CD867174                 570 bp    mRNA    linear   EST 11-JUL-2003
DEFINITION  AZO2.105J01F001124 AZO2 Triticum aestivum cDNA clone AZO2105J01,
            mRNA sequence.
ACCESSION   CD867174
VERSION     CD867174.1  GI:32550990
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 570)
  AUTHORS   Genoplante.
  TITLE     Genoplante, a major partnership french program in plant genomics
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Genoplante
            Genoplante
            93, rue Henri Rochefort 91025 EVRY CEDEX France
            Tel: 33 1 69 47 54 00
            Fax: 33 1 69 47 54 10
            This sequence has been generated in the framework of the french
            plant genomics programme 'Genoplante' (http://www.genoplante.com
            and http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .570
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="recital"
                     /db_xref="taxon:4565"
                     /clone="AZO2105J01"
                     /tissue_type="root"
                     /clone_lib="AZO2"
ORIGIN

  Query Match             8.9%;  Score 99;  DB 6;  Length 570;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-41;
  Matches   99;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1013 GGGGAGGCGGGCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGA 1072
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGGAGGCGGGCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGA 60

Qy       1073 AATCAACAAGCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 1111
             |||||||||||||||||||||||||||||||||||||||||
Db         61 AATCAACAAGCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 99


RESULT 3
CA593541
LOCUS       CA593541                 594 bp    mRNA    linear   EST 21-NOV-2002
DEFINITION  wpa1c.pk002.p19 wpa1c Triticum aestivum cDNA clone wpa1c.pk002.p19
            5' end, mRNA sequence.
ACCESSION   CA593541
VERSION     CA593541.1  GI:25143327
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 594)
  AUTHORS   Tingey,S.V., Moore,G., Griffiths,S., Powell,W., Wolters,P.,
            Dolan,M., Hainey,C., Miao,G., Caraher,N. and Hanafey,M.K.
  TITLE     DuPont Wheat cDNA Sequence in collaboration with the John Innes
            Center 1
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Scott V. Tingey
            Crop Genetics
            E. I. DuPont de Nemours and Company
            1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
            Tel: 302-631-2602
            Fax: 302-631-2607
            Email: Scott.V.Tingey@USA.dupont.com
            Seq primer: M13.
FEATURES             Location/Qualifiers
     source          1. .594
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4565"
                     /clone="wpa1c.pk002.p19"
                     /tissue_type="anthers"
                     /lab_host="DH10B"
                     /clone_lib="wpa1c"
                     /note="Vector: pBluescript SK+; Site_1: EcoRI; Site_2:
                     XhoI; Wheat (Triticum aestivum) pre-meiotic anthers JIC"
ORIGIN

  Query Match             8.3%;  Score 92;  DB 6;  Length 594;
  Best Local Similarity   100.0%;  Pred. No. 8.5e-38;
  Matches   92;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1020 CGGGCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAATCAAC 1079
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGGGCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAATCAAC 60

Qy       1080 AAGCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 1111
              ||||||||||||||||||||||||||||||||
Db         61 AAGCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 92


RESULT 4
CD938039
LOCUS       CD938039                 631 bp    mRNA    linear   EST 15-JUL-2003
DEFINITION  OV.108O07F010206 OV Triticum aestivum cDNA clone OV108O07, mRNA
            sequence.
ACCESSION   CD938039
VERSION     CD938039.1  GI:32785547
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 631)
  AUTHORS   Genoplante.
  TITLE     Genoplante, a major partnership french program in plant genomics
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Genoplante
            Genoplante
            93, rue Henri Rochefort 91025 EVRY CEDEX France
            Tel: 33 1 69 47 54 00
            Fax: 33 1 69 47 54 10
            This sequence has been generated in the framework of the french
            plant genomics programme 'Genoplante' (http://www.genoplante.com
            and http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .631
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="recital"
                     /db_xref="taxon:4565"
                     /clone="OV108O07"
                     /tissue_type="ovary"
                     /clone_lib="OV"
ORIGIN

  Query Match             8.3%;  Score 92;  DB 6;  Length 631;
  Best Local Similarity   100.0%;  Pred. No. 8.5e-38;
  Matches   92;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1020 CGGGCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAATCAAC 1079
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          1 CGGGCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAATCAAC 60

Qy       1080 AAGCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 1111
              ||||||||||||||||||||||||||||||||
Db         61 AAGCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 92


RESULT 5
CD865240
LOCUS       CD865240                624 bp    mRNA    linear   EST 11-JUL-2003
DEFINITION  AZO2.073I23F000912 AZO2 Triticum aestivum cDNA clone AZO2073I23,
            mRNA sequence.
ACCESSION   CD865240
VERSION     CD865240.1  GI:32549056
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 624)
  AUTHORS   Genoplante.
  TITLE     Genoplante, a major partnership french program in plant genomics
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Genoplante
            Genoplante
            93, rue Henri Rochefort 91025 EVRY CEDEX France
            Tel: 33 1 69 47 54 00
            Fax: 33 1 69 47 54 10
            This sequence has been generated in the framework of the french
            plant genomics programme 'Genoplante' (http://www.genoplante.com
            and http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .624
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="recital"
                     /db_xref="taxon:4565"
                     /clone="AZO2073I23"
                     /tissue_type="root"
                     /clone_lib="AZO2"
ORIGIN

  Query Match             8.1%;  Score 90;  DB 6;  Length 624;
  Best Local Similarity   100.0%;  Pred. No. 1e-36;
  Matches   90;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1022 GGCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAATCAACAA 1081
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAATCAACAA 60

Qy       1082 GCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 1111
              ||||||||||||||||||||||||||||||
Db         61 GCGGTGCTTGCCGAGAAGAGAGAGAGAGAG 90


RESULT 6
CA712930
LOCUS       CA712930                430 bp    mRNA    linear   EST 26-NOV-2002
DEFINITION  wdk3c.pk008.e19 wdk3c Triticum aestivum cDNA clone wdk3c.pk008.e19
            5' end, mRNA sequence.
ACCESSION   CA712930
VERSION     CA712930.1  GI:25434723
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 430)
  AUTHORS   Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
            Miao,G., Caraher,N. and Hanafey,M.K.
  TITLE     DuPont Wheat cDNA Sequence
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Scott V. Tingey
            Crop Genetics
            E. I. DuPont de Nemours and Company
            1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
            Tel: 302-631-2602
            Fax: 302-631-2607
            Email: Scott.V.Tingey@USA.dupont.com
            Seq primer: M13.
FEATURES             Location/Qualifiers
     source          1. .430
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4565"
                     /clone="wdk3c.pk008.e19"
                     /tissue_type="kernel"
                     /lab_host="DH10B"
                     /clone_lib="wdk3c"
                     /note="Vector: pBluescript SK+; Site_1: EcoRI; Site_2:
                     XhoI; Wheat (Triticum aestivum L.) developing kernel, 14
                     days after anthesis."
ORIGIN

  Query Match             7.9%;  Score 88;  DB 6;  Length 430;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-35;
  Matches   88;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1024 CCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAATCAACAAGC 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAATCAACAAGC 60

Qy       1084 GGTGCTTGCCGAGAAGAGAGAGAGAGAG 1111
              ||||||||||||||||||||||||||||
Db         61 GGTGCTTGCCGAGAAGAGAGAGAGAGAG 88


RESULT 7
CD884411
LOCUS       CD884411                643 bp    mRNA    linear   EST 14-JUL-2003
DEFINITION  F1.116I20F010507 F1 Triticum aestivum cDNA clone F1116I20, mRNA
            sequence.
ACCESSION   CD884411
VERSION     CD884411.1  GI:32648852
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 643)
  AUTHORS   Genoplante.
  TITLE     Genoplante, a major partnership french program in plant genomics
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Genoplante
            Genoplante
            93, rue Henri Rochefort 91025 EVRY CEDEX France
            Tel: 33 1 69 47 54 00
            Fax: 33 1 69 47 54 10
            This sequence has been generated in the framework of the french
            plant genomics programme 'Genoplante' (http://www.genoplante.com
            and http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .643
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="recital"
                     /db_xref="taxon:4565"
                     /clone="F1116I20"
                     /tissue_type="leaf one"
                     /clone_lib="F1"
ORIGIN

  Query Match             5.9%;  Score 66;  DB 6;  Length 643;
```

```
  Best Local Similarity   100.0%;  Pred. No. 1e-23;
  Matches   66;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1046 CAAGGCAACAACAAAGTGCGCGTGAGAAATCAACAAGCGGTGCTTGCCGAGAAGAGAGAG 1105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CAAGGCAACAACAAAGTGCGCGTGAGAAATCAACAAGCGGTGCTTGCCGAGAAGAGAGAG 60

Qy       1106 AGAGAG 1111
              ||||||
Db         61 AGAGAG 66


RESULT 8
CF132917
LOCUS       CF132917                642 bp    mRNA    linear   EST 24-JUL-2003
DEFINITION  WHE4351_G12_M23ZT Wheat meiotic floret cDNA library Triticum
            aestivum cDNA clone WHE4351_G12_M23, mRNA sequence.
ACCESSION   CF132917
VERSION     CF132917.1  GI:33216504
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 642)
  AUTHORS   Anderson,O.D., Chao,S., Crossman,C., Langridge,P., Lazo,G.R.,
            Pham,J., Rausch,C.J., Sutton,T., Woo,J. and Wilson,C.
  TITLE     The structure and function of the expressed portion of the wheat
            genomes - Meiotic floret cDNA library
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            Sequences have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
            Seq primer: T7 primer.
FEATURES             Location/Qualifiers
     source          1. .642
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="WHE4351_G12_M23"
                     /tissue_type="Whole florets with anthers"
                     /dev_stage="Meiotic stages pre-meiosis-metaphase I"
                     /lab_host="E. coli DH10B"
                     /clone_lib="Wheat meiotic floret cDNA library"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; Plants
                     were grown in a glasshouse. Anther meiotic stage was
                     determined microscopically after removing a single anther
                     from a primary floret. If determined to be between (and
                     including) meiotic stages pre-meiosis and metaphase I,
                     remaining floret tissue was collected and pooled for
                     library construction. The tissue, total RNA, and poly(A)
                     RNA were prepared, cDNA synthesised, and directionally
                     ligated into pSPORT1 by Tim Sutton in the P Langridge Lab
                     at the Department of Plant Science, University of
                     Adelaide, Waite Campus, Australia. Average insert size
                     1.4Kb. Plasmid DNA preparations and DNA sequencing were
                     performed in the OD Anderson lab (all other authors)."
ORIGIN

  Query Match             5.8%;  Score 64;  DB 7;  Length 642;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-22;
  Matches   64;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1048 AGGCAACAACAAAGTGCGCGTGAGAAATCAACAAGCGGTGCTTGCCGAGAAGAGAGAG 1107
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db         47 AGGCAACAACAAAGTGCGCGTGAGAAATCAACAAGCGGTGCTTGCCGAGAAGAGAGAG 106

Qy       1108 AGAG 1111
              ||||
Db        107 AGAG 110


RESULT 9
CD894217
LOCUS       CD894217                712 bp    mRNA    linear   EST 14-JUL-2003
DEFINITION  G118.125L08F010828 G118 Triticum aestivum cDNA clone G118125L08,
            mRNA sequence.
ACCESSION   CD894217
VERSION     CD894217.1  GI:32665532
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 712)
  AUTHORS   Genoplante.
  TITLE     Genoplante, a major partnership french program in plant genomics
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Genoplante
            Genoplante
            93, rue Henri Rochefort 91025 EVRY CEDEX France
            Tel: 33 1 69 47 54 00
            Fax: 33 1 69 47 54 10
            This sequence has been generated in the framework of the french
            plant genomics programme 'Genoplante' (http://www.genoplante.com
            and http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .712
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="recital"
                     /db_xref="taxon:4565"
                     /clone="G118125L08"
                     /tissue_type="grain (118 degrees per day after
                     pollination)"
                     /clone_lib="G118"
ORIGIN

  Query Match             5.8%;  Score 64;  DB 6;  Length 712;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-22;
  Matches   64;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1048 AGGCAACAACAAAGTGCGCGTGAGAAATCAACAAGCGGTGCTTGCCGAGAAGAGAGAG 1107
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         11 AGGCAACAACAAAGTGCGCGTGAGAAATCAACAAGCGGTGCTTGCCGAGAAGAGAGAG 70

Qy       1108 AGAG 1111
              ||||
Db         71 AGAG 74


RESULT 10
CA701748
LOCUS       CA701748                556 bp    mRNA    linear   EST 26-NOV-2002
DEFINITION  wkm2c.pk005.j20 wkm2c Triticum aestivum cDNA clone wkm2c.pk005.j20
            5' end, mRNA sequence.
ACCESSION   CA701748
VERSION     CA701748.1  GI:25423541
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 556)
  AUTHORS   Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
            Miao,G., Caraher,N. and Hanafey,M.K.
```

```
  TITLE     DuPont Wheat cDNA Sequence
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Scott V. Tingey
            Crop Genetics
            E. I. DuPont de Nemours and Company
            1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
            Tel: 302-631-2602
            Fax: 302-631-2607
            Email: Scott.V.Tingey@USA.dupont.com
            Seq primer: M13.
FEATURES             Location/Qualifiers
     source          1. .556
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="hard red spring"
                     /db_xref="taxon:4565"
                     /clone="wkm2c.pk005.j20"
                     /tissue_type="kernel"
                     /lab_host="DH10B"
                     /clone_lib="wkm2c"
                     /note="Site_1: EcoRI; Site_2: XhoI; Wheat (Triticum
                     aestivum L.) kernel malted 175 hours at 4 C"
ORIGIN

  Query Match             5.7%;  Score 63;  DB 6;  Length 556;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-22;
  Matches   63;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1049 GGCAACAACAAAGTGCGCGTGAGAAATCAACAAGCGGTGCTTGCCGAGAAGAGAGAGAGA 1108
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCAACAACAAAGTGCGCGTGAGAAATCAACAAGCGGTGCTTGCCGAGAAGAGAGAGAGA 60

Qy       1109 GAG 1111
             |||
Db         61 GAG 63


RESULT 11
CD904051
LOCUS       CD904051                637 bp    mRNA    linear   EST 14-JUL-2003
DEFINITION  G356.112E12F010920 G356 Triticum aestivum cDNA clone G356112E12,
            mRNA sequence.
ACCESSION   CD904051
VERSION     CD904051.1  GI:32678379
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 637)
  AUTHORS   Genoplante.
  TITLE     Genoplante, a major partnership french program in plant genomics
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Genoplante
            Genoplante
            93, rue Henri Rochefort 91025 EVRY CEDEX France
            Tel: 33 1 69 47 54 00
            Fax: 33 1 69 47 54 10
            This sequence has been generated in the framework of the french
            plant genomics programme 'Genoplante' (http://www.genoplante.com
            and http://genoplante-info.infobiogen.fr).
FEATURES             Location/Qualifiers
     source          1. .637
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="recital"
                     /db_xref="taxon:4565"
                     /clone="G356112E12"
                     /tissue_type="grain (356 degrees per day after
                     pollination)"
                     /clone_lib="G356"
ORIGIN

  Query Match             5.7%;  Score 63;  DB 6;  Length 637;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-22;
  Matches   63;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1049 GGCAACAACAAAGTGCGCGTGAGAAATCAACAAGCGGTGCTTGCCGAGAAGAGAGAGAGA 1108
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 GGCAACAACAAAGTGCGCGTGAGAAATCAACAAGCGGTGCTTGCCGAGAAGAGAGAGAGA 71

Qy       1109 GAG 1111
             |||
Db         72 GAG 74


RESULT 12
BE419251
LOCUS       BE419251                422 bp    mRNA    linear   EST 24-JUL-2000
DEFINITION  WWR021.H3R000101 ITEC WWR Wheat Root Library Triticum aestivum cDNA
            clone WWR021.H3, mRNA sequence.
ACCESSION   BE419251
VERSION     BE419251.1  GI:9417097
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 422)
  AUTHORS   Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T.,
            Cloutier,S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A.,
            Gustafson,P., Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,J.,
            Joudrier,P., Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P.,
            Ogihara,Y., Pecchioni,N., Qualset,C., Schuch,W., Selvaraj,G.,
            Shariflou,M., Sorrells,M., Warburton,M. and Wenzel,G.
  TITLE     International Triticeae EST Cooperative (ITEC): Production of
            Expressed Sequence Tags for Species of the Triticeae
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Schuch W
            Zeneca Wheat Improvement Centre, Norwich Research Park
            Colney Lane, Norwich NR4 7UH UNITED KINGDOM
            Tel: 44 1603 250 2600
            Fax: 44 1603 250 699
            Email: wolfgang.schuch@aguk.zeneca.com
            International Triticeae EST Cooperative (ITEC)
            http://wheat.pw.usda.gov/genome.
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Novosibirskaya 67"
                     /db_xref="taxon:4565"
                     /clone="WWR021.H3"
                     /tissue_type="root"
                     /clone_lib="ITEC WWR Wheat Root Library"
                     /note="M13 Reverse sequencing primer used for 5' end of
                     clone. "
ORIGIN

  Query Match             5.6%;  Score 62;  DB 2;  Length 422;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-21;
  Matches   62;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1050 GCAACAACAAAGTGCGCGTGAGAAATCAACAAGCGGTGCTTGCCGAGAAGAGAGAGAGAG 1109
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          8 GCAACAACAAAGTGCGCGTGAGAAATCAACAAGCGGTGCTTGCCGAGAAGAGAGAGAGAG 67

Qy       1110 AG 1111
              ||
Db         68 AG 69


RESULT 13
BJ257084
```

```
LOCUS       BJ257084                578 bp    mRNA    linear   EST 17-SEP-2002
DEFINITION  BJ257084 Y. Ogihara unpublished cDNA library, Wh_h Triticum
            aestivum cDNA clone whh19m06 5', mRNA sequence.
ACCESSION   BJ257084
VERSION     BJ257084.1  GI:23089660
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 578)
  AUTHORS   Ogihara,Y. and Murai,K.
  TITLE     Expressed genes in Triticum aestivum
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .578
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="whh19m06"
                     /tissue_type="spike at heading date"
                     /dev_stage="Feekes' scale 10.5"
                     /clone_lib="Y. Ogihara unpublished cDNA library, Wh_h"
ORIGIN

  Query Match             5.6%;  Score 62;  DB 4;  Length 578;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-21;
  Matches   62;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1050 GCAACAACAAAGTGCGCGTGAGAAATCAACAAGCGGTGCTTGCCGAGAAGAGAGAGAGAG 1109
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCAACAACAAAGTGCGCGTGAGAAATCAACAAGCGGTGCTTGCCGAGAAGAGAGAGAGAG 60

Qy       1110 AG 1111
             ||
Db         61 AG 62


RESULT 14
BJ221033/c
LOCUS       BJ221033                579 bp    mRNA    linear   EST 17-SEP-2002
DEFINITION  BJ221033 Y. Ogihara unpublished cDNA library, Wh Triticum aestivum
            cDNA clone wh22o13 3', mRNA sequence.
ACCESSION   BJ221033
VERSION     BJ221033.1  GI:23073564
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 579)
  AUTHORS   Ogihara,Y. and Murai,K.
  TITLE     Expressed genes in Triticum aestivum
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .579
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="wh22o13"
                     /tissue_type="spike at meiosis"
                     /dev_stage="Feekes' scale 9"
                     /clone_lib="Y. Ogihara unpublished cDNA library, Wh"
ORIGIN

  Query Match             5.6%;  Score 62;  DB 4;  Length 579;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-21;
  Matches   62;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1050 GCAACAACAAAGTGCGCGTGAGAAATCAACAAGCGGTGCTTGCCGAGAAGAGAGAGAGAG 1109
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 GCAACAACAAAGTGCGCGTGAGAAATCAACAAGCGGTGCTTGCCGAGAAGAGAGAGAGAG 520

Qy       1110 AG 1111
             ||
Db         519 AG 518


RESULT 15
BJ253061/c
LOCUS       BJ253061                597 bp    mRNA    linear   EST 17-SEP-2002
DEFINITION  BJ253061 Y. Ogihara unpublished cDNA library, Wh_f Triticum
            aestivum cDNA clone whf27c17 3', mRNA sequence.
ACCESSION   BJ253061
VERSION     BJ253061.1  GI:23088097
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 597)
  AUTHORS   Ogihara,Y. and Murai,K.
  TITLE     Expressed genes in Triticum aestivum
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .597
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="whf27c17"
                     /tissue_type="spike at flowering date"
                     /dev_stage="Feekes' scale 10.5.1"
                     /clone_lib="Y. Ogihara unpublished cDNA library, Wh_f"
ORIGIN

  Query Match             5.4%;  Score 60;  DB 4;  Length 597;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-20;
  Matches   60;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1052 AACAACAAAGTGCGCGTGAGAAATCAACAAGCGGTGCTTGCCGAGAAGAGAGAGAGAG 1111
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        596 AACAACAAAGTGCGCGTGAGAAATCAACAAGCGGTGCTTGCCGAGAAGAGAGAGAGAG 537


Search completed: August 29, 2005, 06:03:46
Job time : 4264 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 28, 2005, 22:56:20 ; Search time 234 Seconds
(without alignments)
7768.818 Million cell updates/sec

Title: US-09-979-549-2
Perfect score: 1111
Sequence: 1 gaagtcagaaggccgttcag.....ccgagaagagagagagagag 1111

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	39.8	3.6	38503	4	US-09-949-016-17257	Sequence 17257, A
	2	39.8	3.6	39715	4	US-09-949-016-12454	Sequence 12454, A
	3	39.2	3.5	601	4	US-09-949-016-170604	Sequence 170604,
	4	38	3.4	27465	4	US-09-949-016-16561	Sequence 16561, A
c	5	37.6	3.4	832	4	US-09-621-976-2813	Sequence 2813, Ap
c	6	37.2	3.3	601	4	US-09-949-016-31783	Sequence 31783, A
c	7	37.2	3.3	601	4	US-09-949-016-193904	Sequence 193904,
c	8	37	3.3	289	3	US-09-007-005-17	Sequence 17, Appl
c	9	37	3.3	289	3	US-09-244-796-17	Sequence 17, Appl
	10	36.2	3.3	53915	4	US-09-949-016-12221	Sequence 12221, A
	11	36.2	3.3	53915	4	US-09-949-016-12730	Sequence 12730, A
	12	36.2	3.3	53924	4	US-09-949-016-12947	Sequence 12947, A
	13	36.2	3.3	53924	4	US-09-949-016-12948	Sequence 12948, A
	14	35.8	3.2	460	4	US-09-513-999C-32264	Sequence 32264, A
	15	35.2	3.2	1141	4	US-09-806-708B-22	Sequence 22, Appl
	16	35	3.2	27968	4	US-09-949-016-15191	Sequence 15191, A
	17	35	3.2	27968	4	US-09-949-016-15192	Sequence 15192, A
c	18	35	3.2	108310	4	US-09-949-016-16366	Sequence 16366, A
c	19	34.8	3.1	541	4	US-09-270-767-10710	Sequence 10710, A
c	20	34.2	3.1	11740	4	US-09-949-016-13670	Sequence 13670, A
c	21	33.8	3.0	56832	4	US-09-949-016-12976	Sequence 12976, A
	22	33	3.0	168575	3	US-09-426-290-1	Sequence 1, Appli
	23	32.6	2.9	601	4	US-09-949-016-57820	Sequence 57820, A
c	24	32.6	2.9	12603	4	US-09-949-016-17096	Sequence 17096, A
	25	32.6	2.9	20495	4	US-09-949-016-17198	Sequence 17198, A
	26	32.6	2.9	42988	4	US-08-311-731A-128	Sequence 128, App
	27	32.6	2.9	134292	4	US-09-949-016-12158	Sequence 12158, A
	28	32.6	2.9	236474	4	US-09-949-016-13418	Sequence 13418, A
c	29	32.4	2.9	1128	4	US-09-540-236-165	Sequence 165, App
c	30	32.4	2.9	14335	4	US-09-596-002-11	Sequence 11, Appl
	31	32	2.9	7218	1	US-08-232-463-14	Sequence 14, Appl
c	32	31.8	2.9	16044	4	US-09-949-016-16378	Sequence 16378, A
	33	31.8	2.9	48135	4	US-09-949-016-17027	Sequence 17027, A
	34	31.8	2.9	147840	4	US-09-949-016-15236	Sequence 15236, A
	35	31.6	2.8	505	4	US-09-621-976-15639	Sequence 15639, A
c	36	31.6	2.8	38983	4	US-09-949-016-15700	Sequence 15700, A
	37	31.6	2.8	41199	4	US-09-949-016-17269	Sequence 17269, A
	38	31.6	2.8	51252	4	US-09-949-016-16348	Sequence 16348, A
	39	31.6	2.8	84296	4	US-09-949-016-17375	Sequence 17375, A
	40	31.4	2.8	399	4	US-09-621-976-8976	Sequence 8976, Ap
	41	31.4	2.8	1464	4	US-09-583-110-591	Sequence 591, App
	42	31.4	2.8	1476	4	US-09-107-433-2028	Sequence 2028, Ap
	43	31.4	2.8	11443	3	US-08-961-527-49	Sequence 49, Appl
c	44	31.4	2.8	105055	4	US-09-949-016-14001	Sequence 14001, A
	45	31.4	2.8	670689	4	US-09-949-016-12505	Sequence 12505, A

ALIGNMENTS

RESULT 1
US-09-949-016-17257
; Sequence 17257, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17257
; LENGTH: 38503
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(38503)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17257

Query Match 3.6%; Score 39.8; DB 4; Length 38503;
Best Local Similarity 50.8%; Pred. No. 0.14;
Matches 95; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

```
Qy        133 TGGGTGGGCGCGCGAGCTCCCCGAGTTTGAGCCGCAATTTTTTTACATTTTATGGCGATG 192
              | |||| | ||    || ||| |  || ||   ||||||||||   |||| | | ||||
Db      20933 TAGGTGTGAGCCACTGCACCCGGCCTTAGAAAGGCAATTTTTTTTTTTTTTTTTGAGATG 20992

Qy        193 GCGTCAGGCGTTTATCTAGGCGTCTGGGAGGGTACATTTGAAGATGTGCCACCAACTCCA 252
              | ||    ||  || |     ||     |   | || || | | |       |      ||  ||
Db      20993 GAGTTTCGCTCTTGTTGTTGCCCAGGCTGGAGTGCAATGGCACCATCTCAGTTCACCGCA 21052

Qy        253 AACCGACAACCCTGTATCTGAGCATGCCTCATGCCTCTCCTTCATGCCTCCCTTTGGGTG 312
              | |   |   || |  |  |||    ||| ||||||  | ||  |  |  ||  | |
Db      21053 ACCTCTCCCTCCCGGGTTCCAGCGATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTA 21112

Qy        313 AGGTCAT 319
              | |||
Db      21113 CAGGCAT 21119
```

```
RESULT 2
US-09-949-016-12454
; Sequence 12454, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12454
;   LENGTH: 39715
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(39715)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12454

  Query Match             3.6%;  Score 39.8;  DB 4;  Length 39715;
  Best Local Similarity   50.8%;  Pred. No. 0.14;
  Matches   95;  Conservative    0;  Mismatches   92;  Indels    0;  Gaps    0;

Qy        133 TGGGTGGGCGCGCGAGCTCCCCGAGTTTGAGCCGCAATTTTTTTACATTTTATGGCGATG 192
              | |||| | ||    || ||| |  || ||   ||||||||||   |||| | | ||||
Db      20933 TAGGTGTGAGCCACTGCACCCGGCCTTAGAAAGGCAATTTTTTTTTTTTTTTTTGAGATG 20992

Qy        193 GCGTCAGGCGTTTATCTAGGCGTCTGGGAGGGTACATTTGAAGATGTGCCACCAACTCCA 252
              | ||   ||  || |    ||    |  | || || | | |      |      ||  ||
Db      20993 GAGTTTCGCTCTTGTTGTTGCCCAGGCTGGAGTGCAATGGCACCATCTCAGTTCACCGCA 21052

Qy        253 AACCGACAACCCTGTATCTGAGCATGCCTCATGCCTCTCCTTCATGCCTCCCTTTGGGTG 312
              | |   |   || | |   |||    ||| ||||||  | ||  |  |  ||  |  |
Db      21053 ACCTCTCCCTCCCGGGTTCCAGCGATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTA 21112

Qy        313 AGGTCAT 319
              | |||
Db      21113 CAGGCAT 21119


RESULT 3
US-09-949-016-170604
; Sequence 170604, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 170604
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-170604

  Query Match             3.5%;  Score 39.2;  DB 4;  Length 601;
  Best Local Similarity   55.2%;  Pred. No. 0.014;
  Matches   74;  Conservative    1;  Mismatches   59;  Indels    0;  Gaps    0;

Qy        179 ATTTTATGGCGATGGCGTCAGGCGTTTATCTAGGCGTCTGGGAGGGTACATTTGAAGATG 238
              || |  ||| | ||| | ||||| | | |  ||  | || |      | |  ||||||||
Db        222 ATCTGGTGGGGCTGGTGACAGGCCTGTCTGCAGAGGACTTGTCTAAGAAAGATGAAGATG 281

Qy        239 TGCCACCAACTCCAAACCGACAACCCTGTATCTGAGCATGCCTCATGCCTCTCCTTCATG 298
              ||||  |  |    || :|   ||  ||  |||||||   ||  ||   | ||  ||
Db        282 CAACACCCAAACTCCTCCCRCCTGGCTAGATGGGAGCATGGAGCAAACCAACCTTTGCTG 341

Qy        299 CCTCCCTTTGGGTG 312
              ||| |  | |||||
Db        342 CCTGCTGTGGGGTG 355


RESULT 4
US-09-949-016-16561
; Sequence 16561, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16561
;   LENGTH: 27465
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-16561

  Query Match             3.4%;  Score 38;  DB 4;  Length 27465;
  Best Local Similarity   55.2%;  Pred. No. 0.46;
  Matches   74;  Conservative    0;  Mismatches   60;  Indels    0;  Gaps    0;

Qy        179 ATTTTATGGCGATGGCGTCAGGCGTTTATCTAGGCGTCTGGGAGGGTACATTTGAAGATG 238
              || |  ||| | ||| | ||||| | | |  ||  | || |      | |  ||||||||
Db      17411 ATCTGGTGGGGCTGGTGACAGGCCTGTCTGCAGAGGACTTGTCTAAGAAAGATGAAGATG 17470

Qy        239 TGCCACCAACTCCAAACCGACAACCCTGTATCTGAGCATGCCTCATGCCTCTCCTTCATG 298
              ||||  |  |    ||  |   ||  ||  |||||||   ||  ||   | ||  ||
Db      17471 CAACACCCAAACTCCTCCCGCCTGGCTAGATGGGAGCATGGAGCAAACCAACCTTTGCTG 17530

Qy        299 CCTCCCTTTGGGTG 312
              ||| |  | |||||
Db      17531 CCTGCTGTGGGGTG 17544


RESULT 5
US-09-621-976-2813/c
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
```

```
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 2813
;   LENGTH: 832
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 235..399
US-09-621-976-2813

  Query Match             3.4%;  Score 37.6;  DB 4;  Length 832;
  Best Local Similarity   11.6%;  Pred. No. 0.062;
  Matches   40;  Conservative  150;  Mismatches  154;  Indels    0;  Gaps    0;

Qy        540 AATCAGTATGCCTTTCTCTGCATCTTTCTTCATGCAAGCATTAAATACTATAGCTAATCT 599
              ||| | |||   ||  :::  :: ::: ||: :  :: : ::|::: ::::  :  |  :
Db        378 AATAATTATATTTTGTYYWTWWKTWWYWTTYTWTRMWWMKKKARRWYYWWKSTYACASRY 319

Qy        600 ACAGCCAGTTTATTATATAAACAGGCTATATAGCTGACCTGGCAGTGCTATAGAGCCGGC 659
              :     : :: :: :  |:: |     :  :::|   | :|    :  ::|: :::::::
Db        318 RKYTWGWWWYMWKRMMSTRWYCYMCWKCCMYRGRRCAWYTMARGRMWSYAWGKWKSMRSA 259

Qy        660 AGCCGGCTCTTCTATTAGCTTTGCTCTTATGGCTACATCTGTGTGAGCAGTCGATTGATT 719
              :::| ::::::   ::    |   :|:: ::  ::::  | :: ::|   |:  |: :
Db        258 MSMCTRMYYKKGSTYWTMKCTCATWCYWYWKYWKRMWSKTCWSGSRGGYMTSYTSTRSYS 199

Qy        720 CAAACAACAAATCCGGGCGTTCAGCAAGTCGGAATGAATTTCGGCTCATCACTCATTGTC 779
              : :|:::: : :  |:  :|:    : ::  : || ::   :     :|:: ::::   :
Db        198 MYWASWMYTMCWWWGRWWSTYWYMAWGKKWWRYATTWRRAMMWWWAAWTMMWYMWWAWCM 139

Qy        780 GTGGGCCTCACGCGTATTCGCCTAACCGTGTTTGAATCAGACCCTCACGAAGCCACGGCT 839
              : :|  :: :  : : : :   :  : :|:  :  |:  ::::::|:   :: |::  :
Db        138 SSRGAAMYRRTMMMWGYRYWWRKKSYRRTRCAWAYAWKTKRSYYWCWRWKWKRCMMMMMM 79

Qy        840 CCAGCGACCCGTTCACCACGTCAGCCTAAAAAAAGAAAAAAAAA 883
               :|  : ::::        :: |  :::::: :: :  : ::|:
Db         78 AMAYGKTMMMRACWKTRYWRWWAWAMWRMWWTMMMMYYWYWRAM 35


RESULT 6
US-09-949-016-31783/c
; Sequence 31783, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 31783
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-31783

  Query Match             3.3%;  Score 37.2;  DB 4;  Length 601;
  Best Local Similarity   51.9%;  Pred. No. 0.068;
  Matches   84;  Conservative    0;  Mismatches   78;  Indels    0;  Gaps    0;

Qy        133 TGGGTGGGCGCGCGAGCTCCCCGAGTTTGAGCCGCAATTTTTTTACATTTTATGGCGATG 192
              | |||| | ||    || ||| |  || ||   ||||||||||   |||| | | ||||
Db        162 TAGGTGTGAGCCACTGCACCCGGCCTTAGAAAGGCAATTTTTTTTTTTTTTTTTTGAGATG 103

Qy        193 GCGTCAGGCGTTTATCTAGGCGTCTGGGAGGGTACATTTGAAGATGTGCCACCAACTCCA 252
              | ||   ||  || |   ||   |  | || || | | |     |    ||  ||
Db        102 GAGTTTCGCTCTTGTTGTTGCCCAGGCTGGAGTGCAATGGCACCATCTCAGTTCACCGCA 43

Qy        253 AACCGACAACCCTGTATCTGAGCATGCCTCATGCCTCTCCTT 294
              | |   |  || |  |  |||    ||| ||||||  | |
Db         42 ACCTCTCCCTCCCGGGTTCCAGCGATTCTCCTGCCTCAGCCT 1


RESULT 7
US-09-949-016-193904/c
; Sequence 193904, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 193904
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-193904

  Query Match             3.3%;  Score 37.2;  DB 4;  Length 601;
  Best Local Similarity   51.9%;  Pred. No. 0.068;
  Matches   84;  Conservative    0;  Mismatches   78;  Indels    0;  Gaps    0;

Qy        133 TGGGTGGGCGCGCGAGCTCCCCGAGTTTGAGCCGCAATTTTTTTACATTTTATGGCGATG 192
              | |||| | ||    || ||| |  || ||   ||||||||||   |||| | | ||||
Db        162 TAGGTGTGAGCCACTGCACCCGGCCTTAGAAAGGCAATTTTTTTTTTTTTTTTTTGAGATG 103

Qy        193 GCGTCAGGCGTTTATCTAGGCGTCTGGGAGGGTACATTTGAAGATGTGCCACCAACTCCA 252
              | ||   ||  || |   ||   |  | || || | | |     |    ||  ||
Db        102 GAGTTTCGCTCTTGTTGTTGCCCAGGCTGGAGTGCAATGGCACCATCTCAGTTCACCGCA 43

Qy        253 AACCGACAACCCTGTATCTGAGCATGCCTCATGCCTCTCCTT 294
              | |   |  || |  |  |||    ||| ||||||  | |
Db         42 ACCTCTCCCTCCCGGGTTCCAGCGATTCTCCTGCCTCAGCCT 1


RESULT 8
US-09-007-005-17/c
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
;  APPLICANT: Szostak, Jack W.
;  APPLICANT: Roberts, Richard W.
;  APPLICANT: Liu, Rihe
;  TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
;  TITLE OF INVENTION: FUSIONS
;  FILE REFERENCE: 00786/350003
;  CURRENT APPLICATION NUMBER: US/09/007,005B
;  CURRENT FILING DATE: 1998-01-14
;  EARLIER APPLICATION NUMBER: 60/035,963
;  EARLIER FILING DATE: 1997-01-27
;  EARLIER APPLICATION NUMBER: 60/064,491
;  EARLIER FILING DATE: 1997-11-06
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 289
;   TYPE: RNA
```

```
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Translation template
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(289)
;   OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

  Query Match             3.3%;  Score 37;  DB 3;  Length 289;
  Best Local Similarity   10.6%;  Pred. No. 0.049;
  Matches   19;  Conservative   75;  Mismatches   86;  Indels    0;  Gaps    0;

Qy        170 TTTTTTTACATTTTATGGCGATGGCGTCAGGCGTTTATCTAGGCGTCTGGGAGGGTACAT 229
              |||||||   |||||:| |  :     :   :|:|: : :  |: : :|     |: : :
Db        274 TTTTTTTTTTTTTTAYGYCYGYCYAYAYGYAYGYTYTYAYCYGYCYAYGYCYTYGYSYNY 215

Qy        230 TTGAAGATGTGCCACCAACTCCAAACCGACAACCCTGTATCTGAGCATGCCTCATGCCTC 289
              ::    ::: :  ::   :::   ::  :  :: : : : :  :: : :::    :: :
Db        214 NYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNY 155

Qy        290 TCCTTCATGCCTCCCTTTGGGTGAGGTCATGTGCCCTTGGCGGCGAGTGGCTTCCCGTTT 349
              ::: : ::: : ::: :  ::     : : ::: : ::   ::  :  :: : ::: :
Db        154 NYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNY 95


RESULT 9
US-09-244-796-17/c
; Sequence 17, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
;  APPLICANT: Szostak, Jack W.
;  APPLICANT: Roberts, Richard W.
;  APPLICANT: Liu, Rihe
;  TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
;  TITLE OF INVENTION: FUSIONS
;  FILE REFERENCE: 00786/350007
;  CURRENT APPLICATION NUMBER: US/09/244,796
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/035,963
;  EARLIER FILING DATE: 1997-01-27
;  EARLIER APPLICATION NUMBER: 60/064,491
;  EARLIER FILING DATE: 1997-11-06
;  EARLIER APPLICATION NUMBER: 09/007,005
;  EARLIER FILING DATE: 1998-01-14
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 289
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Translation template
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(289)
;   OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17

  Query Match             3.3%;  Score 37;  DB 3;  Length 289;
  Best Local Similarity   10.6%;  Pred. No. 0.049;
  Matches   19;  Conservative   75;  Mismatches   86;  Indels    0;  Gaps    0;

Qy        170 TTTTTTTACATTTTATGGCGATGGCGTCAGGCGTTTATCTAGGCGTCTGGGAGGGTACAT 229
              |||||||   |||||:| |  :     :   :|:|: : :  |: : :|     |: : :
Db        274 TTTTTTTTTTTTTTAYGYCYGYCYAYAYGYAYGYTYTYAYCYGYCYAYGYCYTYGYSYNY 215

Qy        230 TTGAAGATGTGCCACCAACTCCAAACCGACAACCCTGTATCTGAGCATGCCTCATGCCTC 289
              ::    ::: :  ::   :::   ::  :  :: : : : :  :: : :::    :: :
Db        214 NYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNY 155

Qy        290 TCCTTCATGCCTCCCTTTGGGTGAGGTCATGTGCCCTTGGCGGCGAGTGGCTTCCCGTTT 349
              ::: : ::: : ::: :  ::     : : ::: : ::   ::  :  :: : ::: :
Db        154 NYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNY 95


RESULT 10
US-09-949-016-12221
; Sequence 12221, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12221
;   LENGTH: 53915
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(53915)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12221

  Query Match             3.3%;  Score 36.2;  DB 4;  Length 53915;
  Best Local Similarity   57.5%;  Pred. No. 3;
  Matches   65;  Conservative    0;  Mismatches   48;  Indels    0;  Gaps    0;

Qy         963 GCGAGCGCATAAATTCTGATTCCTGCCTGCCTGCCGGACAATTTATCTTTGGGGAGGCGG 1022
               | |||  || | |   ||   |||  ||| |||| | | |  |||||  | ||  | |
Db       44530 GGGAGTCCACAGAAAATGGGGCCTCACTGACTGCTTGTCCAGGTATCTCAGTGGGAGGGA 44589

Qy        1023 GCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAAT 1075
                 ||  | ||| |  |||  | |||||| |||| | ||  |||   |||| |
Db       44590 ATGGGAGTGGGAAATGGAGTACTCAAGGCCACAAGACAGGCCGCCATAGAACT 44642


RESULT 11
US-09-949-016-12730
; Sequence 12730, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12730
;   LENGTH: 53915
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(53915)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12730
```

```
  Query Match             3.3%;  Score 36.2;  DB 4;  Length 53915;
  Best Local Similarity   57.5%;  Pred. No. 3;
  Matches   65;  Conservative    0;  Mismatches   48;  Indels    0;  Gaps    0;

Qy        963 GCGAGCGCATAAATTCTGATTCCTGCCTGCCTGCCGGACAATTTATCTTTGGGGAGGCGG 1022
              | |||  || | |   ||   |||  ||| ||||  | | |  |||||  | ||  | |
Db      44530 GGGAGTCCACAGAAAATGGGGCCTCACTGACTGCTTGTCCAGGTATCTCAGTGGGAGGGA 44589

Qy       1023 GCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAAT 1075
                ||  | ||| |  |||  | |||||| |||| | ||  |||   |||| |
Db      44590 ATGGGAGTGGGAAATGGAGTACTCAAGGCCACAAGACAGGCCGCCATAGAACT 44642


RESULT 12
US-09-949-016-12947
; Sequence 12947, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12947
;   LENGTH: 53924
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(53924)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12947

  Query Match             3.3%;  Score 36.2;  DB 4;  Length 53924;
  Best Local Similarity   57.5%;  Pred. No. 3;
  Matches   65;  Conservative    0;  Mismatches   48;  Indels    0;  Gaps    0;

Qy        963 GCGAGCGCATAAATTCTGATTCCTGCCTGCCTGCCGGACAATTTATCTTTGGGGAGGCGG 1022
              | |||  || | |   ||   |||  ||| ||||  | | |  |||||  | ||  | |
Db      44530 GGGAGTCCACAGAAAATGGGGCCTCACTGACTGCTTGTCCAGGTATCTCAGTGGGAGGGA 44589

Qy       1023 GCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAAT 1075
                ||  | ||| |  |||  | |||||| |||| | ||  |||   |||| |
Db      44590 ATGGGAGTGGGAAATGGAGTACTCAAGGCCACAAGACAGGCCGCCATAGAACT 44642


RESULT 13
US-09-949-016-12948
; Sequence 12948, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12948
;   LENGTH: 53924
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(53924)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12948

  Query Match             3.3%;  Score 36.2;  DB 4;  Length 53924;
  Best Local Similarity   57.5%;  Pred. No. 3;
  Matches   65;  Conservative    0;  Mismatches   48;  Indels    0;  Gaps    0;

Qy        963 GCGAGCGCATAAATTCTGATTCCTGCCTGCCTGCCGGACAATTTATCTTTGGGGAGGCGG 1022
              | |||  || | |   ||   |||  ||| ||||  | | |  |||||  | ||  | |
Db      44530 GGGAGTCCACAGAAAATGGGGCCTCACTGACTGCTTGTCCAGGTATCTCAGTGGGAGGGA 44589

Qy       1023 GCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAAT 1075
                ||  | ||| |  |||  | |||||| |||| | ||  |||   |||| |
Db      44590 ATGGGAGTGGGAAATGGAGTACTCAAGGCCACAAGACAGGCCGCCATAGAACT 44642


RESULT 14
US-09-513-999C-32264
; Sequence 32264, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Duclert, A.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
;  FILE REFERENCE: 59.US2.REG
;  CURRENT APPLICATION NUMBER: US/09/513,999C
;  CURRENT FILING DATE:  2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/122,487
;  PRIOR FILING DATE: 1999-02-26
;  NUMBER OF SEQ ID NOS: 36681
;  SOFTWARE: Patent.pm
; SEQ ID NO 32264
;   LENGTH: 460
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 26
;   OTHER INFORMATION: s=g or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 188
;   OTHER INFORMATION: r=a or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 193
;   OTHER INFORMATION: m=a or c
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 194
;   OTHER INFORMATION: r=a or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 311
;   OTHER INFORMATION: n=a, g, c or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 352
;   OTHER INFORMATION: n=a, g, c or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 353
```

```
;   OTHER INFORMATION: d=a or g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 458
;   OTHER INFORMATION: y=c or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 459
;   OTHER INFORMATION: k=g or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 460
;   OTHER INFORMATION: y=c or t
US-09-513-999C-32264

  Query Match             3.2%;  Score 35.8;  DB 4;  Length 460;
  Best Local Similarity   56.6%;  Pred. No. 0.17;
  Matches   64;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy        963 GCGAGCGCATAAATTCTGATTCCTGCCTGCCTGCCGGACAATTTATCTTTGGGGAGGCGG 1022
              | |||  || | |   ||   |||  ||| ||||  | | |  |||||  | ||  | |
Db        272 GGGAGTCCACAGAAAATGGGGCCTCACTGACTGCTTGTCNAGGTATCTCAGTGGGAGGGA 331

Qy       1023 GCCGGGATTGGAGACAGAGCCCACAAGGCAACAACAAAGTGCGCGTGAGAAAT 1075
                ||  | ||| |  |||    |||||| |||| | ||  |||   |||| |
Db        332 ATGGGAGTGGGAAATGGAGTNDTCAAGGCCACAAGACAGGCCGCCATAGAACT 384


RESULT 15
US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
;  APPLICANT: The University of British Columbia
;  TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
;  FILE REFERENCE: 4810-58741
;  CURRENT APPLICATION NUMBER: US/09/806,708B
;  CURRENT FILING DATE:  2001-04-03
;  PRIOR APPLICATION NUMBER: US 60/147,133
;  PRIOR FILING DATE: 1999-08-04
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
;   LENGTH: 1141
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   NAME/KEY: promoter
;   LOCATION: (1)..(1141)
;   OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAE1 promoters
US-09-806-708B-22

  Query Match             3.2%;  Score 35.2;  DB 4;  Length 1141;
  Best Local Similarity   17.5%;  Pred. No. 0.51;
  Matches   45;  Conservative   79;  Mismatches  133;  Indels    0;  Gaps    0;

Qy        157 GTTTGAGCCGCAATTTTTTTACATTTTATGGCGATGGCGTCAGGCGTTTATCTAGGCGTC 216
              ||           |: ::| ::: ||::::::       :  :| : :::    ::||
Db        833 GTNNNNNNNNNNAYAWWTNKWYYTTDDRWRBAYTNNNNNNRMAYYGAYADDYAYYMSDTC 892

Qy        217 TGGGAGGGTACATTTGAAGATGTGCCACCAACTCCAAACCGACAACCCTGTATCTGAGCA 276
              :     : | :   |   : | ::       :|:    :: | :          || :  :
Db        893 DAWMKWDATKMNNATTYNRGTAWRTNNNNNNMTMKTKYYBHAAWNNNNNNGKMCTAHTWW 952

Qy        277 TGCCTCATGCCTCTCCTTCATGCCTCCCTTTGGGTGAGGTCATGTGCCCTTGGCGGCGAG 336
                  |  |: |:: | ||||::    |  :|:   |:   : ||:::    ||:: :| |
Db        953 VCKATKTTKGCWMNCTTTCRKYKNNCTWYTWMTTTRTTWYAATRWKTNNATGSMTRCNAT 1012

Qy        337 TGGCTTCCCGTTTAGAGCAAGTATAATAAGTCCTAGTCAGCTGGCTATAAGATGTTCCAC 396
               :   :  |::|:   ::: |:::  ::: : | ::   :: :   : :  :|::: ||
Db       1013 GWKNNNYWTGWKTRWTAYRMATRWMKAWWKVMATGSWNTNSYARWAYKTRAYKGWYYNAC 1072

Qy        397 ATCAGCAAATCCTTAAA 413
```

```
              |: ::   |||: |: |
Db       1073 AWRWRWGKATCYMTDNA 1089


Search completed: August 29, 2005, 00:07:34
Job time : 237 secs
```